From fluid flow to coupled processes in fractured rock: recent advances and new frontiers

Hari S Viswanathan¹, Jonathan Blair Ajo-Franklin², Jens Thomas Birkholzer³, James William Carey⁴, Yves Guglielmi⁵, Jeffrey De’Haven Hyman¹, Satish Karra¹, Laura Pyrak-Nolte⁶, Harihar Rajaram⁷, Gowri Srinivasan¹, and Daniel M Tartakovsky⁸

¹Los Alamos National Laboratory (DOE)
²Rice University
³Lawrence Berkeley National Laboratory (DOE)
⁴Los Alamos National Laboratory
⁵Lawrence Berkeley National Laboratory
⁶Purdue University
⁷Johns Hopkins University
⁸Stanford University

November 22, 2022

Abstract

Quantitative prediction of natural and induced phenomena in fractured rock is one of the great challenges in the Earth and Energy Sciences with far-reaching economic and environmental impacts. Fractures occupy a very small volume of a subsurface formation but often dominate flow, transport and mechanical deformation behavior. They play a central role in CO₂ sequestration, nuclear waste disposal, hydrogen storage, geothermal energy production, nuclear nonproliferation, and hydrocarbon extraction. These applications require prediction of fracture-dependent quantities of interest such as CO₂ leakage rate, hydrocarbon production, radionuclide plume migration, and seismicity; to be useful, these predictions must account for uncertainty inherent in subsurface systems. Here, we review recent advances in fractured rock research that cover field- and laboratory-scale experimentation, numerical simulations, and uncertainty quantification. We discuss how these have greatly improved the fundamental understanding of fractures and one’s ability to predict flow and transport in fractured systems. Dedicated field sites provide quantitative measures of fracture flow that can be used to identify dominant coupled processes and to validate models. Laboratory-scale experiments fill critical knowledge gaps by providing direct observations and measurements of fracture geometry and flow under controlled conditions that cannot be obtained in the field. Physics-based simulation of flow and transport provide a bridge in understanding between controlled simple laboratory experiments and the massively complex field-scale fracture systems. Finally, we review the use of machine learning-based emulators to rapidly investigate different fracture property scenarios and to accelerate physics-based models by orders of magnitude to enable uncertainty quantification and near real-time analysis.
From fluid flow to coupled processes in fractured rock: recent advances and new frontiers

H.S. Viswanathan¹, J. Ajo-Franklin²⁻³, J. Birkholzer³, J.W. Carey⁴, Y. Guglielmi³, J.D. Hyman¹, S. Karra¹, L. Pyrak-Nolte⁵⁻⁹⁻¹⁰, H. Rajaram⁶, G. Srinivasan⁷, and D.M. Tartakovsky⁸

¹ Computational Earth Science, Los Alamos National Laboratory, Los Alamos, NM, USA
² Department of Earth, Environmental, and Planetary Science, Rice University, Houston, TX, USA
³ Energy Geoscience Division, Lawrence Berkeley National Laboratory, Berkeley, CA, USA
⁴ Earth System Observations, Los Alamos National Laboratory, Los Alamos, NM, USA
⁵ Department of Physics and Astronomy, Purdue University, West Lafayette, IN, USA
⁶ Department of Environmental Health and Engineering, Johns-Hopkins University, Baltimore, MD, USA
⁷ Physics Division, Los Alamos National Laboratory, Los Alamos, NM, USA
⁸ Department of Energy Resources Engineering, Stanford University, Stanford, CA, USA
⁹ Department of Earth, Atmospheric and Planetary Sciences, Purdue University, West Lafayette, IN, USA
¹⁰ Lyles School of Civil Engineering, Purdue University, West Lafayette, IN, USA

Corresponding author: Hari S. Viswanathan (viswana@lanl.gov)

Key Points:

- Integration of field and lab experiments, simulation and uncertainty quantification is required to understand and predict the properties and response of fracture systems.

- Densely monitored field sites and in situ lab experiments provide quantitative measures of flow and transport that can constrain models.

Abstract

Quantitative prediction of natural and induced phenomena in fractured rock is one of the great challenges in the Earth and Energy Sciences with far-reaching economic and environmental impacts. Fractures occupy a very small volume of a subsurface formation but often dominate flow, transport and mechanical deformation behavior. They play a central role in many subsurface applications including CO₂ sequestration, nuclear waste disposal, hydrogen storage, geothermal energy production, nuclear nonproliferation, and oil and gas extraction. These applications require prediction of fracture-dependent quantities of interest such as CO₂ leakage rate, hydrocarbon production, radionuclide plume migration, and seismicity; to be useful, these predictions must account for uncertainty inherent in subsurface systems such as the number, spatial distribution and connectivity of existing and induced fractures. Here, we review recent advances in fractured rock research that cover field- and laboratory-scale experimentation, numerical simulations, and uncertainty quantification. We discuss how these have greatly improved the fundamental understanding of fracture processes and one’s ability to predict flow and transport in fractured systems. Dedicated field sites provide quantitative measures of fracture flow that can be used to identify dominant coupled processes and to validate both conceptual and numerical models. Laboratory-scale experiments fill critical knowledge gaps by providing direct observations and measurements of fracture geometry and flow under controlled conditions that cannot be obtained in the field. Physics-based simulation of flow and transport provide a bridge in understanding between controlled simple laboratory experiments and the massively complex field-scale fracture systems. Finally, we review the use of machine learning-based emulators and surrogate models to rapidly investigate different fracture property scenarios and to accelerate physics-based models by orders of magnitude to enable uncertainty quantification and near real-time analysis.

Plain Language Summary

Some of the greatest challenges currently facing humanity have roots in the Earth and Energy Sciences. Policy makers are relying on scientific research to answer questions related to the transition to green renewable energy, to the mitigation of the climate crisis, and to ensure global stability with reliable energy and water resources. A common thread in addressing these societal
issues with far-reaching economic and environmental impacts, is the prediction of flow and transport in subsurface systems in the Earth, and in fractured rock in particular. The need to predict, optimize and ultimately control fractured subsurface systems is an increasingly important topic with 80% of the U.S energy resources and 50% of its drinking water supply coming from the subsurface. In this review, we describe the state-of-the-art of research on flow and transport in fracture systems and the path forward for the integration of field observations, laboratory experiments, predictive modeling and uncertainty quantification to enable more efficient and environmentally prudent usage of critical subsurface resources.
1. Introduction

Some of the greatest challenges currently facing humanity have roots in the Earth and Energy Sciences. Policy makers are relying on scientific research to answer questions related to the transitioning to green renewable energy, to the mitigation of the climate crisis, and to ensuring global stability with reliable energy and water resources. A common thread in addressing these societal issues is answering these science questions with far-reaching economic and environmental impacts, is the prediction of flow and transport in subsurface systems in the Earth, and in fractured rock in particular. The need to predict, optimize and ultimately control fractured subsurface systems is an increasingly important topic with 80% of the U.S energy resources and 50% of its drinking water supply coming from the subsurface (Hubbard et al., 2015). Additionally, decision making in the realm of national security, such as nuclear nonproliferation, rely on predictions of gas flow through subsurface fractures to assess the nuclear capabilities of other nation states (Jordan et al., 2014).

Fractures are breaks or mechanical discontinuities in rock that consist of two rough surfaces in partial contact. The voids between contacts provide the flow and transport paths through a fractured rock. Though fractures make up a very small volume of the subsurface, they often are the major conduits that dominate flow and transport behavior in many subsurface environments. In addition, the presence of fractures can strongly impact the geomechanical behavior of a rock mass because fractures are intrinsically planes of weakness, and slight perturbations in stress can open/close the voids that affect flow and transport. Fractured systems are difficult to characterize and predict for several reasons. First, fracture systems are inherently complex because of the geometry and connectivity of a fracture network and of the individual fractures that compose the network. Second, rock is opaque which makes it difficult to identify and image fractures in the subsurface which is required to constrain the properties and topology of fracture networks and the individual fractures within the network. Third, flow and transport properties in fractures are highly sensitive to coupled thermal, hydrologic, mechanical and chemical (T-H-M-C) processes that are triggered by natural and subsurface activities (e.g., Tsang et al., 2012). Finally, any evaluation of coupled processes must consider the contributions from both the fracture network and rock matrix as well as the fluxes between the two. For example, fractures are the interfaces
between rock and the natural and engineered fluids that exist, are injected, extracted or sequestered in the subsurface. However, flow and transport through fractures are moderated by source-sink contributions from the matrix, by chemical alterations that depend on fracture-matrix interactions, and thermal transfer that is mediated by advection into the matrix and affected by the thermal conductivity of the rock matrix (National Research Council, 1996).

These challenges complicate efforts to optimize many key subsurface energy production and anthropogenic waste disposal activities. For example, current extraction methods recover little more than 10% of in-place unconventional hydrocarbons, which is largely believed to be a consequence of sub-optimal fracture performance. Extensive geothermal energy resources remain untapped due to the inability to effectively stimulate and sustain engineered fracture flow. The absence of accurate quantitative predictions of fluid movement and its impact induced seismicity from deep fluid injection into fracture systems has impeded the development of enhanced geothermal energy and the disposal of produced water and CO₂. Lastly, the viability of CO₂ sequestration and high-level nuclear waste disposal remains uncertain because of the difficulty in determining potential flow and transport rates through fractures.

Another challenging aspect of fracture systems arises from the vast disparity in time scales over which fracture processes occur in different subsurface applications. This affects the ability to characterize and model these systems because different processes may dominate at early and late times in the subsurface activity (Figure 1). For example, nuclear nonproliferation monitoring requires detection of gases such as radioactive xenon from underground tests that seeps into the surface within minutes to months through fractures generated and reactivated by complex T-H-M-C processes induced by the nuclear blast. Hydrocarbon extraction from unconventional oil and gas sites operates at time scales of days to years and requires a detailed understanding of hydro-mechanical coupling caused by the hydraulic fracturing processes. Geothermal energy extraction requires full characterization of T-H-M-C processes, with reservoir performance forecasts that range from 1 to 100 years. Disposal of CO₂ and nuclear waste in repositories must account for time scales from days (during injection) to thousands of years to ensure safe containment. At these long time scales, chemical processes often play a critical role in addition to other coupled processes since relatively slow chemical reactions alter the permeability and flow field. As the world transitions from fossil fuels to a carbon-neutral economy, the subsurface
will continue to play an important role. The emerging area of blue hydrogen, where natural gas is used to produce hydrogen and CO\textsubscript{2} is sequestered in the subsurface, is an example of a future energy technology where fractured systems are critical (van Renssen, 2020).

![Figure 1: Approximate relevant timescales for subsurface fracture system applications and principle coupled processes that occur from these subsurface activities.](image)

Analysis and prediction of the behavior of fracture systems also face the challenge of dealing with a wide range of spatial scales while often lacking high-resolution direct measurements. While fluids move through fractures and interact with matrix materials on scales of µm to mm, most subsurface applications affect fracture systems at much larger scales, typically meters to kilometers. At such scales, flow occurs within a network of fractures subject to external stresses and temperature gradients. Useful constraints on the impact of network topology and coupled processes at large scales can be established in underground field experiments (m to 100s of m) (Section 2). Laboratory-scale experiments can be used to characterize individual fractures or small fracture systems, to elucidate the basic hydrologic behavior of fractures, and their response to T-H-M-C processes on the order of 10s nm to 1 m for time scales from ns to months. These controlled simple experiments complement field studies, fill in the knowledge gaps remaining from field observations, and can validate simulations (Section 3).
Small-to-reservoir-scale simulations can integrate the entire system across scales to forecast behavior (Section 4). At all scales, T-H-M-C processes may play a critical role, and physics-based models can interpret observations and help to develop an understanding of the key processes that control flow and transport in fracture systems. However, characterization data are often spatially and temporally sparse such that knowledge of the distribution and properties of fractures is highly limited. Consequently, the ability to equip predictions with uncertainty bounds is crucial. This requires the development of upscaled models and emulators of complex fracture systems that can be used to produce thousands of predictions that span a wide parameter space (Section 5).

Topical reviews addressing the broad spectrum of fracture transport issues are decades old (Tsang, 1991; Wang, 1991; National Research Council, 1996). Some reviews deal with specific technical issues present in fractured systems (Babadagli, 2020; Berkowitz, 2002; Berre et al., 2019; Dietrich et al., 2005; Seales, 2020). Other reviews touch, but do not focus on, flow and transport in fractured rock (Bonnet et al., 2001; Laubach et al., 2019; Meakin and Tartakovsky, 2009; Molz et al., 2004; and Tsang and Neretnieks, 1998). Many of the challenges identified in these works still remain. In particular, the role of coupled T-H-M-C processes has been shown to be even more critical for predicting flow and transport. With intensified interest in fractures, the 2015 National Academy of Sciences report on Characterization, Modeling, Monitoring and Remediation (http://www.nap.edu/21742) has been rereleased to the public in 2020 at no charge. The 2015 DOE roundtable report (Pyrak-Nolte & DePaolo 2015) identified coupled processes as a key need to improve our ability to control fractures and subsurface flow to enhance their resource potential and for safe operation of waste storage and disposal facilities.

High-resolution measurements under extreme conditions are now possible because of the rapid advancements in downhole characterization techniques, the advent of real-time geophysics, and new distributed sensing with fiber optics. These new methods produce “big data” that shifts the data-scarcity paradigm, once very prevalent in subsurface science, to a new paradigm of multiple large yet noisy data sets that need to be combined and properly interpreted to better constrain subsurface simulation. Laboratory experiments that can replicate in situ conditions in real rock have made it possible to measure key parameters under controlled and repeatable conditions to fill knowledge gaps where field observations are lacking. Progress in high-
performance computing has enabled detailed representation of T-H-M-C fracture flow mechanisms that can be constrained by the measurements. The explosion of machine learning has permeated the geosciences facilitating extraction of key parameters from noisy data, discovery of critical mechanisms, and robust uncertainty quantification (Bergen et al., 2019). Finally, dedicated field sites provide quantitative measures of fracture flow processes and our ability to model those. As a result, in the last five years, the potential to integrate subsurface data and simulation has greatly increased in a field that has been plagued by data scarcity. With these advances and the potential to fuse laboratory experiments, field observations and simulation, a new review of the current state of the art is needed.

Our review explores recent advances in new experimental methods, both in the field and in the research laboratory, that provide direct observations of fracture flow; physics-based simulations that span the properties of individual fractures to large complex fracture networks; emulators that can rapidly investigate different fracture property scenarios; and surrogate models that accelerate physics-based models by orders of magnitude to enable uncertainty quantification and near real-time analysis. We discuss how these advances have greatly improved the understanding of, and the ability to predict, fundamental fracture processes and flow and transport in fractured systems. The review is organized as follows:

- **Field Observations:** We describe the recent developments in our ability to monitor the complex perturbations of fractured rock systems as triggered by subsurface applications such as CO$_2$ sequestration and geothermal energy extraction. These perturbations, which may involve coupling between thermal, hydrologic, geomechanical and geochemical processes, in turn may lead to transient changes in the flow and transport behavior of fracture systems. We discuss the truly transformational advances made recently in our ability to develop and deploy dynamic monitoring techniques, providing measurements that capture the temporal variation of a property or system state at sufficient spatial resolution. We describe the value of dedicated *in situ* field test sites for subsurface research which allow for testing of new monitoring techniques and provide testbeds for new subsurface utilization ideas involving fractured systems. Two examples are presented of ongoing experiments conducted in underground research facilities that bring
together state-of-the-art monitoring technologies, fast data processing, and integrated modeling to better understand dynamic subsurface processes in fractured rock.

- **Laboratory Experiments:** We discuss advances in understanding the effect of fundamental chemical and mechanical processes on geometry and flow. Recent laboratory experiments replicate subsurface pressures, temperatures, and stresses under controlled conditions and provide data through direct optical, x-ray and acoustic observations of fluid flow and transport in fractured rocks. These results provide single fracture-level information on the interaction of rock properties, stress, displacement and chemistry on fracture aperture and flow.

- **Coupled Thermal, Hydrologic, Mechanical and Chemical Simulations in Fractured Rocks:** Field and lab experiments illuminate the effect of coupled processes that often control flow and transport in subsurface fractured systems. We describe here how a new generation of mechanistic flow and transport fracture models built on discrete and continuum approaches uses high performance computing to predict T-H-M-C flow and transport in highly complex fracture networks and how these predictions can be validated by this latest generation of lab and field experiments.

- **Upscaling, Emulators and Surrogate Models:** We describe here how advances in emulators and surrogate models using machine learning and graph theory make it possible to accelerate mechanistic models of flow and transport in fractures by three to four orders of magnitude. This enables both analysis with robust uncertainty quantification to bound system behavior and near real-time analysis, as many applications now require the ability to optimize performance of subsurface fractured reservoirs to increase production efficiency and reduce environmental impacts. We also discuss upscaling methods to best incorporate small scale fracture flow processes into larger scales.

2. **Field Observations and Experiments**

A first step in the understanding and prediction of coupled processes that control flow and transport in fractured rock at the field scale involves the identification and characterization of relevant natural fractures, in particular those that are hydraulically significant. A second step
involves monitoring the complex perturbations of fractured rock systems caused by subsurface applications such as CO₂ sequestration, geothermal energy extraction, or geothermal energy extraction. These perturbations may lead to transient changes in the flow and transport behavior of the fractured system, which in turn are induced by complex coupling between thermal, hydrological, geomechanical and geochemical processes. In other words, the first step involves the static characterization of fractured rock hydrology (via geophysical methods, well imaging and logging, and hydrologic/tracer testing; see National Research Council, 1996), while the second comprises dynamic monitoring of relevant changes to the behavior of fractured rocks. Both steps are challenging, and both have seen important developments over the past decades. However, truly transformational advances have recently been made in our ability to develop and deploy dynamic monitoring techniques, providing measurements that capture the temporal variation of a property or system state at sufficient spatial resolution. Such techniques have allowed for high-resolution monitoring of the temporal and spatial evolution of complex processes and perturbations, and these advances will be our focus. We provide an overview of how the advent of real-time geophysics and integrated borehole tools for coupled processes, coupled with high-resolution spatial monitoring via fiber optics technologies, enabled by much improved methods for rapid synthesis, interpretation and data analysis, and demonstrated in dedicated in situ research facilities, has resulted in much better observation of fractured rock perturbations and how these affect flow and transport processes.

2.1. Advances and developments for monitoring complex perturbations in fractured rock systems

2.1.1. Dynamic monitoring techniques

Dynamic monitoring techniques have dramatically increased the availability of high resolution in-well data in fractured systems. These techniques can be schematically classified into permanent sensors, mainly fiber optic based, grouted behind casing, and non-permanent probes that can perform local tests as well as more or less continuous profiles by circulating along the borehole. Distributed fiber-optic sensing (DFOS) methods, a family of techniques which utilize the analysis of scattered laser light within fiber strands (Hartog et al., 2017), are capable of making measurements of temperature, strain, and even seismic wavefields at fine spatial intervals (mm to 10s of m) and as a function of time; these techniques can integrate the complex
fracture-intact rock interactions along multi-kilometer borehole axes, thus providing a unique combination of spatial extent and spatio-temporal resolution during monitoring. Distributed temperature sensing (DTS), of both passive and active forms, has been particularly valuable in understanding flow in fractured systems (e.g., Read et al., 2013; Coleman et al., 2015; Maldaner et al., 2019). More recently, distributed acoustic sensing (DAS) has been leveraged for low frequency hydro-mechanical measurements in crystalline rocks (e.g., Becker et al., 2017). In the context of unconventional reservoirs, strain measured by DAS has quickly revolutionized detection of fracture propagation to offset wells, crucial for understanding induced fracture architecture after stimulation (e.g., Jin and Roy, 2017); this has been enabled by the low noise floor (sub nanostrain) of modern DAS interrogators which will only improve with advances in optical technology and engineered fibers. Higher frequency recordings using the same DAS systems can also be utilized for microseismic event location (e.g. Verdon et al., 2020), relevant for understanding the stimulated rock volume, as well as active source seismic imaging to detect elastic perturbations induced by fractures (e.g. Binder et al., 2020; Titov et al., 2021).

A current limitation of DFOS methods, both for strain and seismic monitoring, is that they are single component measurements. While multi-fiber cables, theoretically capable of resolving the full strain tensor, have been proposed and modeled (Lim et al., 2018), such systems are yet to be realized in operational packages. DSS monitoring often tends to display a relatively poor sensitivity to quasi-static shear, resulting in a limited measurement resolution of any environmental loading that is not parallel to the borehole axis (although some DFOS settings such as helical fibers attempt to improve on that). DFOS methods also require careful fiber encapsulation and protection during deployment to avoid damage during casing and completion, particularly in deep deviated wells.

In parallel to distributed sensing, the most recent borehole logging tools offer a full three-dimensional (3D) coverage of the borehole surface, i.e., axial, azimuthal and radial measurements (see Serra and Serra 2004 for a synthesis). Image-log techniques such as Electrical MicroImager (EMI), Fullbore Formation MicroImager (FMI), a wide variety of acoustic scanning tools, and finite difference time domain radars (FDTD) allow the local reconstruction of Discrete Fracture Models (DFN) when the tool is profiling along the borehole nearfield (i.e.,
about 1-2m away from the borehole). When combined with tracer tests, such tools can then be used for time-lapse monitoring of electrically conductive tracers penetrating fractures, and fracture mechanical opening (see Aghli et al., 2020 for case study). Finally, since the 2000s (Murdoch et al., 2003), hydromechanical well testing has been developed, using probes that allow the local downhole direct measurement of fractures opening and shear while pressurizing it between two sealing inflatable packers (Schweisinger et al., 2009; Guglielmi et al., 2014). These techniques provide direct input for geomechanical analyses of fracture properties, full stress tensor and potential for induced seismicity (Guglielmi et al., 2020b).

2.1.2. Data synthesis, interpretation, inversion, and analysis

Dynamic high-resolution measurement techniques have led to such drastic increases in data quantity that the incoming data streams are too large to be stored, transmitted, or analyzed within reasonable time periods. To solve this issue in the domain of borehole log analyses, a wide variety of statistical approaches (clustering, multivariate analyses, artificial intelligence algorithms) has been developed to extend the classical fracture detection, density, geometry and various attributes analyses from image logs (Massiot et al., 2017; Zhang et al., 2011; Zazoum, 2013; Dong et al., 2020). These techniques enable advanced corrections from the inherent sampling bias related to borehole logging and provide probability laws for building different near borehole DFNs equiprobable scenarios. Fully coupled thermo-hydro-mechanical (T-H-M) numerical codes such as TOUGH-FLAC (Rutqvist et al., 2003), FLAC3d and 3DEC (Itasca Consulting Group, 2013), and many other academic codes which have been continuously developed in parallel to the in situ characterization techniques provide the theoretical basis for the forward interpretation of fracture permeability variations measured during hydromechanical well tests (Murdoch and Germanovich, 2006; Cappa et al., 2006b). Recently, Afshari Moein et al. (2018) developed a stress-based tomography approach that combines a geomechanical modeling of quasi-static stress variations caused by fractures with an inversion algorithm to reconstruct probable DFN geometries from stress variability profiles obtained from borehole logs. In addition, progress has been made on inversion methods, mainly through the combined inversion of different borehole log datasets (Tian et al., 2021) or by using some borehole datasets to better constrain the inversion of other types of cross-borehole datasets (Wu et al., 2019).
Geophysical monitoring strategies for fractured systems have also benefited from a range of new techniques over the last decade, including both physics-based and data-driven strategies to extract even more relevant information from geophysical measurements. A major advance in seismic imaging, predominantly driven by advancement of computational resources, has been the wide adoption of full wavefield inversion (FWI) strategies as well as wave-based reverse-time migration (RTM) methods. These techniques, while developed theoretically and first implemented in the 1990s (e.g., Tarantola 1988, Pratt, 1998, Pratt et al., 1999) have now become feasible for inversion of 3D volumes (Virieux and Operto, 2009; Brittan et al., 2013), because of the availability of massively parallel computation and more recently GPU and cloud deployments. The strength of these approaches is their capacity to utilize a larger fraction of the seismic waveform to recover elastic properties with higher spatial resolutions than previously possible with arrival-time based methods. This is a particularly crucial element when imaging fractured systems where localized variations in seismic anisotropy or attenuation are the target feature (e.g., Bretaud et al., 2014) or a detailed map of the host rock is needed to guarantee the integrity of engineered systems (e.g., Bentham et al., 2018). Most recently, the use of FWI approaches in a timelapse capacity has opened a window into monitoring subtle property changes (e.g., Li et al., 2020) relevant to understanding complex fractured system behavior.

For systems involving engineered fractures including those generated by hydraulic stimulation in both geothermal and shale production, perhaps the most useful geophysical monitoring tool is the measurement of microseismic events to allow mapping of zones where new flow paths are being created, or at least where rocks are being actively deformed. While the location of small seismic events has a long history, recent processing advances have allowed for real-time hypocenter, magnitude, and in some cases focal mechanism estimation, an impressive suite of parameters useful for operational guidance. Even in cases involving massive data volumes acquired at either high recording rates or high sensor densities, the use of edge processing allows for close to real-time locations using conventional (e.g., Schoenball et al., 2020) or fiber-optic (Verdon et al., 2020) sensing arrays. Real-time microseismic analysis for fracture mapping is also being facilitated by the use of machine learning approaches to accurately automate laborious components of the processing flow including first arrival picking (Chai et al., 2020).
2.1.3. Dedicated field test sites

Unlike other scientific fields where the bulk of the research can be explored in laboratory experiments, it is critically important to extend fractured rock research to the field scale. In field settings, scientific understanding can be gained, and technology innovations can be tested under in situ conditions, across compartments and scales, in natural geologic environments with their complexities and heterogeneities, and in the presence of natural driving forces. Particularly valuable are heavily characterized and densely monitored sites dedicated to and available for long-term research activities. While the characterization and monitoring tools deployed at these sites might not be feasible or cost-efficient in industrial contexts, the resulting highly resolved datasets improve our understanding of the relevant physical principles, aid in development of conceptual models and the testing of predictive simulation approaches (Birkholzer et al., 2018, 2019), allow for field-testing of new monitoring techniques, and not least provide testbeds for new subsurface utilization ideas involving fractured systems. While forty years ago there were only a few dedicated in situ test sites with a focus on fractured rock behavior and feasibility of nuclear waste disposal (International Atomic Energy Agency, 2001), we have seen a few dozen of such facilities be established in the past two decades, in a variety of geological environments and covering a broad range of subsurface applications. Often utilized in open team science settings, with researchers from multiple organizations (and sometimes industry stakeholders) working together, these subsurface testbeds have advanced frontier applications in fractured rock understanding and utilization, spanning fracture hydrology research and applications such as geothermal and fossil energy, CO₂ storage, energy storage, and radioactive waste disposal.

Below, we will step through time and provide some selected case studies of the scientific advances made in dedicated field test sites.

In the 1980’s, emphasis was placed on flow and transport processes in fractured systems, often in low permeability hard rock environments, which were then a preferred target for developing geological repositories for radioactive waste. As fracture zones are the dominant pathways for radionuclide transport in such systems, in situ studies focused on (1) identifying and characterizing permeable fracture zones and (2) measuring and predicting the distribution of
groundwater flow and solute transport. The Stripa Mine in Sweden is an early example of a community testbed to investigate the behavior of granitic rocks from the standpoint of their potential use as a repository for radioactive waste isolation (Witherspoon, 2000; Witherspoon et al., 1980). A Swedish-American cooperation launched in the late 1970s which later included multiple other partners, the project used existing mine tunnels several hundred meters below the ground surface to launch one of the most comprehensive studies of flow processes through a large volume of low permeability granitic rock (150 m by 150 m by 50 m). Detailed fracture mapping was conducted in mine tunnels and in many boreholes drilled from the drifts. Cross-hole and single-hole radar measurements were made to determine the orientation and extent of fracture zones at the site. Seismic techniques were used successfully to determine the orientation and extent of fracture zones. Multiple logs were run, borehole deviation, sonic velocity, single-point resistance, normal resistivity, caliper, temperature, borehole fluid conductivity, natural gamma radiation, and neutron porosity. Further, a hydraulic testing program involved pressure measurements, single-hole and cross-hole flow testing in identified fracture zones, large-scale permeability testing based on inflow into the tunnels, and finally tracer testing. Together, these studies provided not only integrated dataset that allowed constructing a reliable conceptual model of fracture flow paths for a site (including a refined representation of their heterogeneity), but also demonstrated the suitability of radar and seismic techniques to correctly describe the geometry of relevant fracture zones (National Research Council, 1996).

The research program and collaborative emphasis at the Stripa Mine served as a blueprint for other “community” approaches to \textit{in situ} research in underground research laboratories (NEA, 2013; Bettini, 2014). For example, in 1995 the Äspö Hard Rock Laboratory opened in Sweden extending almost 500 m deep into granitic rock (Bäckblom et al., 1997), followed a year later by the Mont Terri Rock Laboratory in an argillite formation in Switzerland (Bossart et al., 2017). Both are purpose-built research facilities with hundreds of meters of drift and alcove space for an impressive number of \textit{in situ} experiments (the research program at Mont Terri consisted of 138 individual experiments between 1996 and 2016), both have initially focused on research associated with radioactive waste isolation but later broadened in scope to other geosciences investigations, and both have an organizational structure that enables extensive collaboration when it comes to sharing technological expertise and experiences. In both cases, an initial focus
on local geologic and hydrogeologic characterization was followed by experiments on repository-induced perturbations, such as the hydro-mechanical effects from tunnel construction and ventilation or from thermally-induced mechanical and chemical processes stemming from the emplacement of heat-emanating radioactive waste. Field testing thus shifted from “static” observation to “dynamic” monitoring of transient changes, using emerging technologies that allow for high-resolution spatio-temporal sensing in harsh environments. Prominent in situ test examples on perturbations of natural rocks include heater experiments like the Prototype Repository at Äspö (Johannesson et al., 2007) or the Full-scale Emplacement Demonstration at Mont Terri (Müller et al., 2017), both of which are full-scale replicas of future waste emplacement tunnels with heating elements representing the waste canisters.

The idea of dedicated, heavily monitored in situ test sites for subsurface research was also pursued in several borehole-based experiments. In the United States, the federal government recently launched an ambitious program of “field laboratory” test sites for unconventional hydrocarbon production, often in close collaboration between industry-driven operations and research institutions. A prominent example is the “Hydraulic Fracturing Test Site” (HFTS) in the Midland Basin in Texas, where a detailed characterization and monitoring strategy was designed to elucidate the hydraulic fracturing and stress interaction behavior in 11 parallel horizontal wells, during stimulation and production. In addition to a full set of geophysical measurements and various other observations of reservoir behavior before, during, and after the completion of about 400+ fracture stages in 2015, the project featured a unique post-stimulation core well drilled through the newly stimulated fractured rock volume. The obtained cores provided a rare opportunity for field studies to be directly informed by detailed laboratory studies involving (1) direct characterization of fracture and proppant distribution, (2) determination of the potential modification of the rock matrix adjacent to hydraulic fractures, and (3) state-of-the-art micro-scale characterization and testing (Courtier et al., 2017; Ciezobka et al., 2018; Birkholzer et al., 2021). A similar field laboratory program was established recently in the United States to advance geothermal energy production in hot dry rocks, which need fracture stimulation to allow for fluid circulation and energy harvesting: The Frontier Observatory for Research in Geothermal Energy (FORGE) is a dedicated community research site in Milford, Utah (Moore et al., 2020) where scientists and engineers are developing, testing, and accelerating breakthroughs in
enhanced geothermal systems (EGS) technologies and techniques. Together, the emergence of such heavily characterized and densely monitored sites, either borehole based or in deep underground research laboratories, has not only led to the successful deployment of new monitoring technologies but has also advanced our understanding of the complex coupling between thermal, hydrologic, geomechanical and geochemical processes in perturbed fractured rock systems.

2.2. Two case studies

Here we present two ongoing field studies which are ground-breaking in their objective of bringing together next-generation monitoring technologies, fast data processing, and integrated modeling to better understand dynamic subsurface processes in fractured rock.

2.2.1. EGS Collab

The EGS Collab project strives to improve our understanding of creating EGS through densely monitored mesoscale stimulation experiments in a crystalline rock mass approximately 1.5 km deep at the Sanford Underground Research Facility (SURF), South Dakota, USA (Figure 2a), formerly known as the Homestake Gold Mine (Kneafsey et al., 2020). A dedicated testbed consisting of eight subhorizontal boreholes of 60-m length was designed to study the creation and function of a subsurface heat exchanger based on hydraulic fracturing to connect an injection-production borehole doublet (Schoenball et al., 2020). To monitor the coupled mechanical, thermal, and hydrogeologic processes occurring during stimulation, the boreholes were instrumented with a multimodal instrument string that included a fiber optic cable for distributed sensing of temperature (DTS), strain (DSS), and acoustic (DAS) signals; electrode strings for electrical resistivity tomography (ERT); and thermistors and piezoelectric seismic sources for continuous active seismic source monitoring (CASSM) (Daley et al., 2007), hydrophones, and three-component accelerometers. The borehole locations were identified using laser survey mapping of the borehole wellheads in the drift and gyro log surveys of the borehole trajectories. Automated workflows based on high network connectivity, virtual team collaboration software, excellent remote sensor display, cloud data exchange, and edge processing allowed for rapid processing and visualization of monitoring data to the degree that experimental controls could be
handled remotely in near-real time (Kneafsey et al., 2020). The first suite of experiments at EGS Collab successfully demonstrated the creation of a fracture network connecting the injection and production wells.

Important lessons were learned through the deployment of multi-modal complementary monitoring tools and coupled multi-physics simulation, such as that the stimulated volume consists of multistrand hydraulic fractures as well as reactivated pre-existing structures, and that the hydraulic fracture growth is strongly influenced by rock fabric, pre-existing fractures, and stress heterogeneities. Planning for another suite of experiments is well under way in a nearby testbed where the stress conditions will allow for hydro-shearing rather than hydraulic fracturing as the prime EGS stimulation method.

2.2.2. Fault Slip (FS) experiments in Mont Terri

The Fault Slip (FS) experiments in the Mont Terri underground research laboratory involve high-resolution monitoring of controlled injection into a highly deformed low-permeability fault zone. The purpose of FS is to examine the dynamic hydromechanical relationship between fluid pressure buildup due to injection, fault deformation, induced seismicity, and creation of permeable pathways along a shale fault. Better understanding of fault activation in a shale is important in the context of geological carbon storage as aseismic and seismic events may jeopardize the integrity of a sealing caprock overlying a CO₂ sequestration reservoir. The test design involves an injection borehole drilled from the overlying Mont Terri tunnel system which intersects the entire fault zone. The section of the injection borehole that intersects the fault is isolated by two inflatable packers and contains a high-resolution probe (the SIMFIP probe) that allows the continuous monitoring of 3D displacements of the fault synchronously with injection pressure and flowrate. A typical activation experiment consists of increasing the pressure step-by-step in the injection zone until fault slip is triggered. When fault is at slip, pressure is maintained for several minutes to study the relationships between slip, opening, permeability variations and induced micro-seismicity. A series of controlled-injection experiments conducted in 2015 allowed the observation of mostly aseismic fault rupture growth from the fluid injection point in the fault to a monitoring point located a few meters away along the fault (Jeanne et al.,
Significant pressure increases were required for the injected fluid to migrate through the initially almost impermeable and mechanically weak fault planes connected to the injection source. Guglielmi et al. (2020a) proposed (but could not directly confirm by measurements) a fault slip mechanism that starts with localized dilatant deformation and channelized flow along initially “closed” but heterogeneous fault zones and is eventually followed by shear failure when a sufficiently large area is pressurized. In 2020, a new suite of experiments was installed in a nearby fault testbed, this time adding advanced high-resolution techniques to be able to track the propagating rupture and pressurized fault patches and their heterogeneities (referred to as FS-B, see Figure 2b). The ongoing experimental campaign envisions multiple injection tests (including with CO$_2$-brine mixtures) over the next few years and comprises new types of measurements to track the long-term leakage in the fault, including (1) time-lapse active seismic imaging across the fault zone using boreholes that straddle the fault from above and below, and (2) DFOS of acoustic signals, temperature and strain. In addition, the new test series hopes to develop and deploy, for the first time, a fiber-optics based sensing of fluid geochemistry in a fault with propagating CO$_2$-brine mixtures. Much will be learned from this unique combination of advanced downhole deformation probes and distributed active and passive sensing, about how high-pressure fluids force their way along a shale fault zone, whether leakage caused by induced seismicity occurs in a channelized and/or heterogeneous manner, if the spatiotemporal characteristics of evolving leakage patterns are correlated (or not) with measurable signals of fault movements, and last not least if water chemistry plays a role in changing the fault’s hydromechanical behavior.
2.3. What have we learned

This section provides an arguably subjective list of high-level learnings achieved by applying advanced monitoring technologies (Section 2.2) in valuable field experiments, like those described in Section 2.3.

2.3.1. Hydromechanical well testing helped to refine constitutive relationships and processes at the single fracture scale by integrating the effects of the in situ fracture rock environment.
Cappa et al. (2008) highlighted how the poroelastic response of a single fracture embedded in a fracture network depends on both the fracture intrinsic properties (stiffness and hydraulic aperture) and on shear strain distribution within its environment. Indeed, compared to laboratory scale tests, this introduces significant non-linearity in the response. More precisely, Murdoch and Germanovich (2012) showed that due to such mechanical interactions with its environment, the flowing fracture compliance varies in both space and time, thus affecting its storage capacity. When natural faults or fractures are brought to rupture by internally increasing fluid pressure, Guglielmi et al. (2020a and b) observed a complex interplay between fault slip and fault opening in low permeability shale formations, with most of the hydraulic conductivity variation in faults caused by opening rather than slip. Moreover, Guglielmi et al. (2015) showed that a large part of fault rupture and associated fluid leakage was aseismic, and thus could not be evidenced by measuring seismic events. These results from field testing at the Mont Terri underground research laboratory showed how important it is to consider the hydromechanical coupled processes to best constrain fault rupture propagation and arrest related to fluid diffusion.

### 2.3.2. Distributed fiber optic sensing enabled high-resolution probing of active fractures and their thermo-hydromechanical interactions within a fracture network.

Becker et al. (2020) showed that distributed sensing with fiber optics can probe nano strain oscillations across flowing fractures along a borehole drilled through a fracture network, thus opening avenues to identify such flowing fractures in a fractured reservoir. Zhang et al. (2020) used coupled forward hydromechanical modeling to show how low frequency DAS and DSS signals can inform on interactions between flowing and non-flowing fractures within a fracture network. Such analyses could enable “imaging” the aseismic fracture flow effects, and thus strongly complement the use of classical seismic monitoring. Munn et al. (2020) used DTS fibers to isolate non-connected, naturally connected and “forced” connected (i.e., during a pumping test) fractures within a fractured aquifer, which is valuable information for many types of subsurface uses. Recently, Hopp et al. (2021) showed that DSS can be used to track the localized deformation of the Mont Terri shale fault (Section 2.2.2) in response to a gallery excavation 30–50 m away. In a fault injection experiment performed at the same site in 2021, magnitudes of DSS, DAS and local SIMFIP measurements set in boreholes crosscutting the fault zone were
found to be highly consistent with each other and complementary. Indeed, the DAS and DSS showed that deformation progressively spread across the entire thickness of the fault zone as a result of high-pressure fluid injection. These preliminary results are crucial in terms of better understanding fault activation by fluid injection, and whether this activation may occur as localized and fast slip or rather a more diffuse and slow displacement.

2.3.3. Multi-modal monitoring, joint inversion and integrative forward modeling have demonstrated the value of complementary data to provide additional constraints on interpretation.

In the EGS Collab experiment described above, Schoenball et al. (2020) used multimodal geophysical measurements—including borehole logs, DSS-DTS-DAS fibers cemented in boreholes, fluid pressures and induced seismicity—to reconstruct a 3D network of fractures hydraulically activated during a cross-hole fluid injection between two boreholes spaced about 20 meters. Combining multi-tracer tests with the analysis of other geophysical data, Neupane et al. (2020) showed a complex correlation between changes in tracer recoveries, shifting of major production from one well to the others, and the repeated hydraulic stimulations conducted in the EGS-Collab experiment. Among changes that most affected the tracer recovery were fast changes in well pressure and the cooling effect related to the injection of water colder than the formation. These were discussed as direct consequences of complex thermo-hydromechanical interplays between the stimulated fractures. Such multimodal approaches proved to be very useful to “image” complex interactions between rock fabric, pre-existing fractures, and stress heterogeneities significantly away from the borehole nearfield, thus opening perspectives to expand from the inter-borehole scale to monitoring at the reservoir scale.

2.4. New frontiers and future work

Recent developments in monitoring strategies for fractured systems have significantly advanced our understanding of flow, transport, and coupled processes in fractured rock. However, challenges remain and more breakthroughs are needed, both in terms of technological advances for field testing and in terms of elevating our scientific understanding of fractured system processes. Below is a selected list of remaining challenges related to both technological approaches and scientific questions.
2.4.1. Technology advances

Current cutting-edge technologies, i.e., 3D downhole probes and distributed multimodal optical fibers, offer the possibility to track signals over spatial scales from nano to micro-scale localized all the way distributed over kilometer deep boreholes while potentially spanning large time scales, potentially annual to decadal for well-supported permanent instrumentation. This new ability offers a promising perspective of probing long-duration phenomena and their effects on fracture flow, such as creep or sealing processes. Advanced instruments that measure strain or strain rate at very high sensitivities may enable the detection of aseismic processes and thus may complement wave-based and active geophysical techniques. Nevertheless, fully accessing 3D data in the field still remains a challenge for certain properties or processes. For example, direct monitoring of shear displacements would lead to better constraints on the tensorial nature of many parameters such as stress, fractured rock elasticity and permeability. Also geophysical monitoring of chemical changes in distant fractures remains elusive. The development of a distributed chemical fiber-optic based sensor (DCS) is a promising path to complement existing DTS-DSS-DAS fiber deployments, particularly in reactive systems where dissolution/precipitation reactions strongly couple to permeability. A related challenge is to develop instruments that can (i) probe deep high-temperature fractured rocks and (ii) are designed to withstand years to decades of field deployment in harsh environments. Also needed are better techniques for joint assessment and inversion of multiple parameters at the same location, such as combining seismic waves, strain measurements, pore pressure, and fluid chemistry data. This would facilitate a better understanding of the changes in fracture hydrogeology caused by remote loadings from seismic waves triggered by distant earthquakes, from stress transfers in stimulated reservoir wells as well as from earth tides.

Meanwhile, new developments to improve the spatial and time-lapse resolution of flow and strain images in fractured rocks will result in the generation of increasingly larger datasets that will require significantly improved data transmission methods (for example using 5G) and processing capabilities. Already, technologies such as DAS can generate upwards of 10s of TB/day. Advanced machine learning techniques are already available for data reduction and analysis; with further improvements, these may offer a possibility for “ultra-fast” big data
analysis and interpretation, enabling both new science breakthroughs and adaptive near real-time control of subsurface operations.

2.4.2. Science advances

The direct multimodal measurement and observation of the chemical behavior of fractures and faults remains very challenging at the field scale. This is especially true in hot fractured geothermal reservoirs, where the interplay of chemical reactions with flow, transport, heat transfer, and mechanics is particularly important while borehole instruments that can tolerate very high temperatures and harsh chemical environments are rare or still in development. Using advanced T-H-M-C numerical modeling, Taron and Ellsworth (2010) highlighted the complex reaction of fractured rocks to the strong stress field rearrangements caused by the injection of cold water, potentially leading to widespread shear failure and significant induced seismicity. They also showed that a significant part of stress dissipation could be related to strong chemo-mechanical fluid rock interaction, causing chemical compaction and aseismic creep of hot hydrated fractures. Such couplings and their relevance have clearly been demonstrated at the laboratory scale: For example, Polak et al. (2003) identified competing effects of decreasing fracture aperture with increasing rock temperature and chemical precipitation (Polak et al., 2003), Luo et al. (2017) demonstrated that fracture hydraulic conductivity increases with a reduction in fluid viscosity, and Andreani et al. (2008) showed that cyclic flows of CO$_2$-brine and CO$_2$-gas can cause an abrupt fracture aperture increase controlled by decohesion of the clay framework induced by CO$_2$-gas acidification. A targeted effort is needed to conduct field studies with a focus on chemo-mechanical behavior, at dedicated field observatories (such as at the FORGE site described above) where new breakthrough technologies for conducting and monitoring the EGS systems are deployed or in the fault testbed at the Mont Terri site where new mesoscale fault activation experiments, with a focus on the exploration of complex H-M and C coupling during and following fault slip, are in preparation.

A key issue for many subsurface applications such as industrial-scale CO$_2$ sequestration or permanent disposal of nuclear waste is to understand and to be able to predict the long-term (years, decades, centuries) hydromechanical evolution in fractures and faults that play a
significant role in the upper crust with regards to elastic strain release and hydrogeological evolution. However, there is a lack of direct observations of fracture closure/healing/sealing, aseismic deformation, slow movement and deformation of fractures over time. For example, the long-term hydrologic evolution of faulted caprocks is poorly constrained although basin-scale seismic observations indirectly indicate the existence of frequent active or ‘fossil’ focused fluid escape paths, some of them aligned with faults or fractured zones (Cartwright & Santamarina, 2015). The study of long-term strain release mechanisms at scale, such as slow slip and creep, is important to improve our understanding of how such mechanisms can modulate the frequency and magnitudes of natural and induced earthquakes as observed from indirect surface geodetic data (Wdowinski, 2009).

The ultimate goal is to achieve the same ultra-dense distributed and time-lapse sensing demonstrated at the tens-of-meter scale in the EGD Collab and Fault Slip experiments (Section 1.2.1 and 1.2.2) in full-scale subsurface applications, such as reservoir-caprock systems affected by pressure buildup from geological CO₂ sequestration at scale or deep basement structures perturbed by stress transfers from overlying fluid manipulations. Dense time-lapse sensor arrays would enable the interrogation of the role of fractures at the crustal scale to provide, potentially, answers to these and other complex research questions. For example, fractures are very sensitive to stress in the deep crystalline basement, a key parameter driving their hydro mechanical behavior. This effect may be less pronounced at shallower depths, where contrasts in the hydromechanical properties of different sedimentary layers may play a more significant role. We have observed that the faults and fractures in the Mont Terri caprock analog behave differently than in the SURF basement analog, which may be caused by many factors including mineralogical differences, different fault architectures, and different constitutive laws describing the T-H-M coupled processes leading to fault permeability variations. How all of these complex processes interact with stress at the crustal scale remains relatively poorly observed. An interesting example is the observed escape of gas through chimneys in many sedimentary basins that develop over time from deep reservoir faults to vertical conduits that outflow at the surface (Løseth et al., 2011). Another example is the effect of very small pressure changes on deep basement faults that are kilometers away from an injection site but are able to activate large magnitude earthquakes (Keranen and Weingarten, 2018). These two examples highlight the
necessity to distribute high-resolution measurements across multiple different depths at relevant scales, to improve our understanding of how T-H-M-C processes drive complex fracture behavior as a function of depth, and how processes acting at a given crustal depth can influence processes occurring at shallower or deeper depths.

3. Laboratory Experiments

Field measurements, described in the previous section, provide fundamental constraints on the geometry and effective transport properties of fracture systems. However, these observations are not well-suited for the investigation of fundamental mechanisms governing transport processes that are necessary to develop a model of fracture systems. In particular, the role of coupled processes such as the impact of stress on fracture permeability or chemical reactions on solute transport are difficult to isolate based on observations of the collective behavior measured in field tests. Laboratory experiments provide a means of controlling key variables and conditions to analyze coupled processes in a reproducible manner. This informs interpretations of field observations and, importantly, provides a mechanism for validating numerical simulations of fracture processes. The simulations, in turn, help overcome the relatively small scale of experiments, as validated simulations can be used to upscale laboratory results to larger-scale, field problems.

The basic conceptual framework for experimental studies is that an individual fracture is the fundamental unit controlling flow and transport through a fracture network and it provides baseline information as input to models used to interpret or predict field observations. In this section, we describe experimental work conducted in the past several decades that has contributed to our understanding of the single fracture base unit and is organized around two fundamental hypotheses. First, knowledge of the geometric properties of a fracture (roughness, aperture, contact area) is sufficient to predict flow and transport properties. Second, geophysical measurements at the field scale that can characterize the state or changed states of fractures can be used to infer changes in some aspects of fracture transport properties.

These hypotheses have led to the following experimental objectives:
• Prediction of flow and transport properties of fractures from a geometric description of fracture properties (roughness, aperture, contact area)
• Determine how coupled processes (stress, displacement, reaction, etc.) modify transport processes by modifying fracture geometry
• Develop methods to characterize and create fractures under more realistic, subsurface conditions that reflect fracture properties as they may exist in the subsurface (in situ x-ray computed tomography, geophysical techniques, microfluidics, etc.).

Below, we describe the major advances that have been achieved in imaging and characterization technology including profilometry, x-ray computed tomography, geophysical techniques, and optical visualization methods in microfluidics and transparent analogs that enable pursuit of these objectives. Similar to field-scale tests described in section 2, laboratory experiments have benefited from improvements in computer hardware, software, data storage and sensor development that enable the use of more sensors and imaging devices to collect more data over a wide range of length and time scales at higher sampling rates during multimodal measurements. In laboratory experiments, fracture geometry and the evolution of this geometry can now be measured with unprecedented detail in real time with x-ray and optical methods. These data sets have led to the discovery of dimensional scaling relationships for roughness and to a lesser extent aperture that drive simulations of transport in fractures. These observations have also shown that transport processes in nature cannot be understood without accounting for the multistranded, non-planar, network character of fracture systems or the individual fractures within the system, and for the fracture intersections that control fracture connectivity. The past decades have yielded an improved understanding of the effect of normal stress on fracture geometry and fluid flow, which has led to the development of a near universal scaling relationship between fluid flow and fracture specific stiffness which can be potentially monitored with geophysical techniques. Coupled processes involving chemical dissolution are well studied, while effects of relative fracture surface displacements are coming into focus but the impacts of precipitation reactions are largely unknown. We find that despite significant progress we lack a definitive prediction of flow and transport based on geometry and that a cohesive framework for interpreting fracture permeability measurements across different fractures in different rock types is lacking. These are important obstacles to the development of a more predictive understanding
of coupled processes. We close with an outlook on opportunities for model validation and future experimental work.

3.1. Recent significant advances and developments in fracture characterization

Laboratory studies of fracture properties and transport behavior have advanced in four key areas: 1) high-resolution profilometry to characterize, *ex situ*, fracture surface roughness; 2) x-ray tomography that characterizes, *in situ*, fracture geometry and that can image the dynamic evolution of fracture systems; 3) optical visualization of coupled flow and reaction processes including multiphase phenomena in microfluidics and transparent analogs; and 4) elastic wave interrogation of fractures for the characterization, prediction and monitoring of flow coupled with stress and reaction, which is a technology that can potentially be field deployable. These techniques have enabled much higher fidelity investigation of the relationship between fracture geometry and flow, which is crucial to fracture simulation; have opened unprecedented views into the 4D evolution of fracture geometry, have provided crucial data for understanding coupled processes; and have demonstrated that elastic waves reveal fracture stiffness-fracture flow relations, which sets that foundation for remote characterization of fracture systems. In this section, we describe the development of these methods and their applications to the study of fractures.

3.1.1. High-resolution profilometry: surface roughness characterization

The void geometry and contact area of a fracture arise from two rough surfaces in partial contact and is the primary link between the hydraulic and mechanical properties of a fracture. The starting measurement for many laboratory studies is to quantify fracture surface roughness to reconstruct the void aperture. In early studies, surface roughness was measured by recording the vertical motion of a stylus, repeatedly performing this measurement across the surface in a grid-like fashion (Brown & Scholz 1985; Swan 1983; Tarantola 1986). The technique was slow: 12 days were needed to move a stylus over 0.3 m² fracture surface in 25 µm increments (Keller and Bonner 1985). A disadvantage of the stylus-based profilometry is the potential of damaging the surface. This led to the use of more rapid, non-contacting methods that use white light or lasers instead of a stylus to probe a surface with µm-scale vertical resolution. Based on Ameli et al.
(2013), a 0.3 m² fracture could be scanned with a lateral spatial resolution of 20 µm in less than 20 minutes.

Several other modern approaches to roughness characterization include LIDAR (light detection and ranging), a laser-light based method more often used in the field and on large surfaces (Candela et al., 2009; Mastrorocco et al., 2018; Polyear and Fairley, 2011; Wilson et al., 2011); high-speed photography, digital holographic microscopy; interferometric microscopy; optical coherence tomography and diffraction phase microscopy. LIDAR has been used to examine the evolution of fault-surface roughness as a function of slip (Sagy et al., 2007), the link between the aspect of ratio of faults and strength (Brodsky et al., 2016) and to construct discrete fracture networks from outcrops for flow simulations (Wilson et al., 2011). Field measurements of surface roughness using LIDAR combined with noncontact, profilometry methods have been used to determine the self-affine fractal scaling behavior of roughness across 9 orders of magnitude (Candela et al., 2012) and to develop quantitative relationships between surface roughness and slip distributions (e.g., Angheluta et al., 2011; Nigon et al., 2017). Other examples include the use of digital holographic microscopy to measure in real-time nano-scale dissolution of gypsum surfaces over a roughly 220 x 220 µm² region (Feng at a., 2017).

3.1.2. High-resolution CT characterization of fracture geometry

In contrast to profilometry techniques, computed tomography (CT) provides a means of directly visualizing intact fracture systems. CT allows direct visualization of fracture apertures in their natural state without piecing the two sides of the fracture together, but at the cost of a generally lower resolution of fracture roughness. High-fidelity analysis of fracture apertures requires optimizing CT image quality (Cnudde and Boone 2013; Ketcham and Carlson 2001; Noiriel 2015; Wildenschild and Sheppard 2016; Withers et al., 2021) as well as the use of fracture-specific CT segmentation strategies (Deng et al., 2016; Huo et al., 2016; Johns et al., 1993; Lee et al., 2021; Ketcham et al., 2010; Voorn et al., 2013). One of the key innovations is the use of “missing attenuation” methods in which sub-resolution characterization of apertures can be achieved by assigning gradations in the x-ray attenuation between matrix and void to fractional quantities of fracture space (Ketcham et al., 2010).
Early measurements generally employed medical x-ray CT with limited resolution (1000s of µm) but were able to measure the statistical distribution of apertures and non-uniform (fingering) of an injected tracer (Johns et al., 1993). With time, instrumentation has improved through the application of industrial scanners (10s of µm), micro-focus CTs (1s of µm), and high-flux and synchrotron sources (1s of µm to 10s of nm). In general, higher resolution comes at the cost of reduced specimen sizes: industrial CT works with several cm diameter specimens and synchrotron studies at a resolution of 1-7 µm operate on specimens of 3-9 mm (e.g., McBeck et al., 2019; Renard et al., 2018; Voltolini and Ajo-Franklin 2019; Godinho et al., 2019).

The potential power and advantage of CT methods are that they yield a fracture volume that is a relatively low-resolution but faithful representation of fluid pathways. This includes a map of the fracture contact area, but resolution limitations make it impossible to prove that “closed regions” are truly closed. Other methods to determine contact areas and to enhance CT analysis include use of injected Wood’s metal that resolves apertures down to 2 µm (Pyrak-Nolte et al., 1997) and use of pressure sensitive film (e.g., Nemoto et al., 2009). CT methods lend themselves to producing time-sequence images that permit investigation of changes to fracture geometry as a function of displacement, compression, and chemical reactions (McBeck et al., 2020c; Renard et al., 2016; Voltolini and Ajo-Franklin 2020; Crandall et al., 2017; Frash et al., 2017, 2019; Menefee et al., 2020). A combination of high-resolution profilometry with aperture-defining CT provides a nearly complete representation of fracture geometry, lacking only a well-resolved definition of contact area.

3.1.3. Transparent analog systems and microfluidics

While x-ray CT can track the progress of fluids through a fracture system (Arshadi et al., 2018; Hirono et al., 2003; Landry et al., 2014; Karpyn et al., 2007), it often lacks the resolution (temporally and spatially) to capture details of multiphase behavior, chemical reactions, and flow fields. Transparent analog systems and microfluidics provide an alternative and essentially synthetic approach to studying flow in fracture systems. In this technique, flow through channels or on the surface of a material is observed by light transmission methods or optical or
fluorescence microscopy yielding µm-scale spatial resolution and sub-second temporal resolution through digital imaging. Realistic surfaces can be created, repeatedly, by etching, machining, texturing or 3D printing methods that mimic fracture cross-sections or surfaces. Transparent analogs and microfluidics systems are constructed of acrylic, glass, silicon, reactive compounds, rocks, minerals and 3D printing. For more direct comparisons with geological materials, micromodels can be constructed of minerals, rocks or other geomaterials (Porter et al., 2015; Mohammadi and Mahani 2020; Gerami et al., 2017; Song et al., 2012), operated at high pressure and temperature conditions (Porter et al., 2015; Morais et al., 2020), or operated with normal stress (Detwiler 2008). The wetting properties of these materials can be modified by chemical treatment (Grate et al., 2012). These have been used to assess a wide variety of phenomena including the impact of dissolution and stress (Detwiler 2008); dissolution/precipitation under multiphase flow conditions (Jiménez-Martínez et al., 2020); turbulent flow and chemical reactions (Lee and Kang 2020); gravity-driven homogeneous precipitation in a fracture (Xu, 2020), and multiphase flow (Keller et al., 1997). 3D printing has enabled fabrication of repeatable structural and material properties to test specific hypotheses to provide a mechanistic understanding of different aspects of coupled processes. 3D printed samples have been used to explore the role of layering and mineral fabric orientation on fracture surface roughness (Jiang et al., 2020), to create porous media with reactive properties to mimic rock (Anjikar et al., 2020); to test hypotheses on how compaction and dissolution of rock microstructure affect transport (Head & Vanorio, 2016) and to measure and simulate the fluid flow properties of fractured porous media for either simulated fracture networks or reconstructed from X-ray computed tomographic images (Martinez et al., 2017; Suzuki et al., 2019). Both analog and microfluidic systems hold great promise for elucidating flow, transport, fracture-matrix interactions and chemical reaction phenomena that may need to be considered in mechanistic models.

3.1.4. Elastic wave characterization of fracture properties

Elastic wave or seismo-acoustic measurements provide a means of transferring laboratory measurements to field monitoring methods that are not possible with the previously discussed techniques. The objective of these studies is to detect fractures and to use changes in the elastic
response of fracture systems as a signal of changes in fracture geometry that control the hydraulic properties of a fracture (Pyrak-Nolte & Nolte, 2016). In the 1980s and 1990s, theoretical developments (Schoenberg, 1980) and experimental measurements (Pyrak-Nolte et al., 1990) of compressional and shear wave transmission across single fractures demonstrated that elastic wave interaction with fractures depends on the frequency of the wave, the matrix material and fracture specific stiffness. Fracture specific stiffness is an effective parameter that captures the complicated geometry of two rough surfaces in contact and depends on both the spatial and probabilistic distributions of contact area and aperture (Hopkins, 1991). For any given frequency, some subset of discontinuities will be optimal for detection because different wavelengths sample different subsets of fractures (Worthington, 2007; Pyrak-Nolte, 2019). The lessons learned from monitoring fractures in the laboratory are conceptually transferrable to the field.

During the past decades, laboratory characterization of fractures has seen a rise in scanning and an increase in the number of sensors employed along a line or 2D region (Acosta-Colon et al., 2009; Li et al., 2009). However, future major advances in laboratory ultrasonic methods for fracture characterization and monitoring will rely on (1) phased arrays composed of independent, miniature ultrasonic sensors and (2) transportable mini-sources (McNab and Campbell 1987). The advantage of using a phased array sensor for probing a fracture is the ability to quickly generate and receive signals from multiple incident angles. How a signal scatters as a function of angle of incidence is sensitive to the internal fracture geometry and the external dimensions (Paris et al., 2006) of a fracture. As shown by Nakagawa et al. (2004), interpretation of the scattered wave-field from rough fracture can provide a map of local fracture stiffness.

Transportable mini-sources generate signals from inside a fracture and have the potential to determine fracture flow path connectivity. Acoustic emitters based on micro-electromechanical systems (MEMs) can be produced at mm-scale (Rassenfoss 2016) for use in larger fracture systems, but have significant potential for jamming in the narrowest aperture along a flow path. For laboratory applications, this challenge has been overcome by using a dissolvable transportable source referred to as chattering dust (Pyrak-Nolte et al., 2020). A dust grain composed of sucrose contains numerous bubbles of compressed gas that bursts concussively as
the dust dissolves. If the dust encounters a constriction, it dissolves until it is small enough to fit through the constriction. During this time, it continues to emit and provides an indication of a change in aperture. Using chattering dust at the field scale will require research into how to extend dust lifetime, development of ultra-low energy sensors and signal discrimination software (Pyrak-Nolte et al., 2020). The combination of these geophysical methods with direct observation techniques described above will provide a powerful approach to analyzing results from field studies described in Section 2.

3.2. What we learned

The new fracture-flow characterization methods described above have given a much more detailed picture of how fracture geometry impacts flow which has brought us closer to achieving Objective 1 (in Section 3.1), a predictive geometry-flow mode. However, as described below we have not yet resolved all of the questions. Research has increasingly focused on more realistic fractures (in terms of fracture types and stress conditions) that bring greater insight into fracture-permeability variation. However, progress is hampered by a lack of a standard that allows inter-comparison of measurements from different experiments on different fractures in different rock types. Our understanding of coupled processes (Objective 2) has greatly improved, particularly for the impact of normal stress and dissolution on flow path geometry in single fractures. Some headway on fracture displacement has been achieved but relatively little is known about impact of precipitate on fracture geometry. Combining multiple geophysical techniques (e.g., acoustic and electrical see Sawayama (2019)) shows promise in the remote characterization of fracture geometry and fracture network permeability (Objective 3). This section reviews these advances and discusses key challenges for research on coupled processes.

3.2.1. Hydrologic properties of fractures

*The cubic law, fracture roughness, fracture aperture, and fracture contact area*

The basic hypothesis is that a full, predictive understanding of fracture flow requires a complete geometric description of a fracture that involves surface roughness, fracture aperture, and
fracture contact area. Aperture by itself is insufficient as while it defines the available flow geometry the resolution is generally insufficient to describe the impact of roughness where the fracture surfaces are close together. Roughness by itself is insufficient as it does not account for the impact of the separation of the surfaces. Contact area is necessary because the mean aperture is meaningless if the flow channel is locally blocked by contacting surfaces, and the spatial distribution of contact area will determine if there is single or multiple flow paths through the fracture void geometry. Most experimental approaches have focused on detailed analysis of the impact of one or two of these key factors. Unfortunately, simultaneous measurements of all three factors are rare to non-existent. Here we briefly review basic concepts and discuss some of the recent results and questions they raise. (Reviews in this general area during the past 20 years include those of Auradou 2009; Berkowitz 2002; Zhang and Chai 2020).

The basic conceptual framework for fracture flow is the cubic law that shows the intimate connection between aperture and roughness: for example, Witherspoon et al.’s (1980) $k = \frac{b^2}{12f}$, where $k$ is the fracture permeability, $b$ is the mean fracture aperture, and $f$ is a factor accounting for surface roughness. Note that the traditional cubic law does not account for contact area, although modifications have been proposed that include a contact area term (see summaries of the various forms of cubic law in Zimmerman and Bovardsson 1996 and Zhang and Chai 2020). The most rigorous attempts to test the cubic law have applied it at the resolution of the measurements (local cubic law). However, even these results overestimate measured fracture permeability (Nicholl et al., 1999; Wang et al., 2015).

Application of the cubic law rests on the assumption that flow is 2D in the plane of the fracture and that flow is laminar. General guidelines as to whether flow in fractures satisfies these conditions are elusive. Flow in 2D requires that the magnitude of roughness variations must be smaller than the characteristic wavelength of the roughness (Zimmerman et al., 1991) and that the Reynolds number be much less than 1. As with the 2D requirement, laminar flow contains a characteristic length scale that is not readily defined or measured. Detailed fluid dynamic simulations in 3D illustrate how roughness within a generally laminar flow system can create non-linear flow patterns such as eddies (Cardenas et al., 2007; Dang et al., 2019; Wang et al., 2016; Zhou et al., 2018). These features impact transport properties through the creation of tails
in breakthrough distributions, as shown by Cardenas et al. (2007) using a tomography-based numerical geometry and by Stoll et al. (2019) in experiments on naturally fractured granite.

While fracture roughness is conceptually simple, studies of the roughness of natural surfaces (see Profilometry section) find that roughness is described by self-affine fractals at scales from microns to kilometers (Candela et al., 2012; Power and Tullis 1991; Schmittbuhl et al., 1995). Self-affine statistics indicate that it is not possible to characterize roughness by a single parameter (e.g., standard deviation of surface height) as roughness grows with the scale of observation (Brown and Scholz 1985). While the change in roughness with spatial dimension is relatively consistent among different fracture systems, the magnitude of the roughness is particular to the fractures in question. Converting surface roughness data into a fracture aperture distribution is fraught with inaccuracies that can arise during alignment - an alignment can only be achieved to within the resolution of the surface profilometry method.

There is far less data on fracture apertures. Caliper-based, field measurements show that the maximum apertures of fractures (both shear and tensile) scale linearly with fracture length (Scholz 2010). However, the mean aperture is not uniform across a fracture with the maximum located near the fracture mid-point with diminishing apertures toward the fracture tips (Vermilye and Scholz 1995). In contrast, direct measurements of small-scale fractures in the lab using x-ray tomography (Section 3) result in Gaussian or lognormal, or bimodal aperture distributions that do not generally resolve differences in aperture between fracture mid-points and tips (Bertels et al., 2001; Crandall et al., 2017; Johns et al., 1993; Keller 1998; Ramandi et al., 2017; Stoll et al., 2019; Wenning et al., 2019). Thus apertures, like roughness increase with scale but unlike roughness, apertures vary systematically across the fracture face.

Contact area plays a critical role in defining the connectivity and number of flow paths in a single fracture. For example, Ishibashi et al. (2018) used numerical simulations to show a non-linear decrease in permeability with increasing contact area. Contact area also is important in transmitting stress across a fracture that results in deformation of both the asperities in contact and the adjacent matrix, which can significantly modify fracture flow geometry (Pyrak-Nolte et
Because the exact local distribution of contact area governs flow paths, a generalized relationship between contact area and permeability remains elusive.

We do not yet have a definitive demonstration of the validity of the local cubic law in reproducing experimental data. There is a critical need for studies that include measurements of roughness, aperture, and contact area. X-ray CT methods would appear to have the most promise for this task but often too low resolution to obtain surface roughness data. Gouze et al. (2003) studied dissolution of limestone using XCT with a resolution of 5 µm that enabled simultaneous characterization of both aperture and roughness. Interestingly, they observed “non-topographic surfaces” development in an impure limestone where surface topography was undercut by dissolution. Song et al. (2021) found consistency between surface roughness profilometry and XCT data in a study of fracture permeability as a function of confining stress. They found that measured hydraulic and mechanical apertures were related by the XCT-measured fracture contact area without consideration of roughness.

In summary, fracture permeability is a complex function of roughness, aperture, and contact area distributions, spatial and probabilistic. Each of these is subject to scale effects and each dynamically evolves with changes in normal stress and fracture displacement. We need to determine what is the necessary resolution of these measurements and what is the minimum amount of information needed to describe or simulate a fracture system to be within say 10-20% of the experimental value. New experiments with full geometric characterization are needed that allow permeability measurements free of edge effects (stress or otherwise).

Permeability of fractures

One of the most important questions in field (Section 2) and numerical studies (Section 4) is the permeability of subsurface fractures. There are now many studies of fracture permeability on a wide range of rock types, under a wide range of environmental conditions, and often involving various coupled processes. Unfortunately, we lack a basis for comparing these measurement values so that we can begin to isolate and understand the impacts of rock type/mineralogy, stress, and displacement on fracture permeability. For a purely elastic rock, the impact of mineralogy,
stress and displacement are linked through the moduli of the rock, but fluid flow through a fracture is only implicitly linked to these parameters through the deformed fracture geometry. The deformed geometry is captured by fracture specific stiffness which is an effective parameter but this has only been studied for purely elastic rock, and not all experiments measure fracture stiffness. Presumably, a comparison of fracture permeabilities would be based on fracture geometry as discussed in the previous section but also will need to include information or a measurement of the fracture deformational properties for different rock types (e.g. ductile versus brittle rock). As a start, we organize the varying results in terms of the type and character of fracture investigated.

Permeability studies have been conducted on fractures that are artificial (saw-cut), induced in the laboratory by tensile (representing rock joints and hydraulic fractures) or shear stress (representing tectonic faults and reactivated fractures), or obtained from natural specimens having tensile or shear characteristics.

*Saw-cut fractures.* These have the disadvantage of being highly idealized, planar features but have the advantage of having controllable surface properties that permit repeatable property investigations. In many cases, the surfaces are polished with a specified grit giving a uniform level of roughness (e.g., Noriel and Deng 2018) or the surfaces can be milled to give structured roughness (e.g., Acosta et al., 2020). These studies are particularly useful in comparing different rock types and mineralogy (Fang et al., 2018). However, many flow studies prop the surfaces apart so no contact area exists between the two surfaces.

*Induced tensile fractures.* These are induced by splitting an unconfined cylindrical sample (e.g., Vogler et al., 2018). This is a natural type of fracture but because the two halves of the specimen separate, the fracture must be reassembled to define an aperture. Hydraulic fracturing at triaxial or true-triaxial conditions can also be used to allow *in situ* study (without disturbing pressure/temperature conditions) of tensile fracture permeability (e.g., Fraser-Harris et al., 2020). Induced tensile fractures tend to generate well-mated fracture surfaces.
*Induced shear fractures.* Modified direct-shear methods can be used to create shear fractures that allow study of *in situ* fracture-permeability relations (e.g., Carey et al., 2015; Frash et al., 2016, 2017, 2019; Shokouhi et al., 2020; Wang et al., 2021). Shearing results in loss of matedness between the surfaces from asperity damage.

*Natural fractures.* These include tensile and shear fractures as “made by nature” (e.g., Crandall et al., 2017; Stoll et al., 2019). However, these are not without artifacts as the original stress conditions have been released (possibly permanently changing fracture aperture or displacement). Sometimes natural fractures have mineral coatings of various thickness that are easily altered over time and affect the repeatability of the measurements.

*Fracture gouge.* Studies of the permeability of mineral powders or natural gouge powders sandwiched between rock or other rigid materials simulate mature fracture systems (e.g., Bakker et al., 2016). As expected, the permeability will be strongly affected by the amount of compaction of gouge and the irreversible nature of the deformation of the gouge porosity in response to stress.

*Synthetic fracture materials and replicas.* These include mineral powders sintered to a rock-like state (Fang et al., 2018), 3D printed materials that allow control of surface roughness and other properties (Fang et al., 2018; Ishibashi et al., 2020; Liu et al., 2021), casts of actual fractures (Hakami and Lasson 1996), or designed fractures surfaces (Detwiler 2008).

Adding to the complexity, permeability studies have been conducted on a wide range of rock types and synthetic materials. Permeability has been measured at a range of temperatures, stress conditions, and fluid pressures and many experiments involve additional parameters and coupled processes (e.g., chemical reactions or stress changes). A method of predicting (or even ballpark estimates) subsurface fracture permeability based on knowledge of rock type, fracture type, fracture history, and stress conditions remains elusive.

*Real vs. idealized fractures and fractures in the subsurface*
Fractures have additional complexity beyond departures from the cubic law. There are at least three considerations complicating models that represent fractures as two individual surfaces in contact: 1) a distinction between mature and immature fractures; 2) non-planar fracture geometries; and 3) the multistranded or fracture network character of real fracture systems. Mature fractures are characterized by substantial displacement that evolved to establish an extensive, well-defined surface containing significant gouge or intervening rock fragments and material. Any flow and transport that occurs between such surfaces is occurring within fragmented material with flow paths more akin to porous media.

Immature slip systems, on the other hand, have much less displacement, lack gouge (but can have fragmented material), have nonplanar segments, and are better described as a series of partially interconnected shear and dilation features, e.g., as seen in en échelon fracture systems (Aydin and Schultz 1990; Segall and Pollard 1980, 1983). In such systems, flow is either anastomosing around obstacles or transferring amongst a complex network of fracture segments. Recent experimental work has attempted to replicate such juvenile fracture formation (Frash et al., 2017, 2019; Shokouhi et al., 2020). These experiments create shear fractures systems and measure permeability without disturbing the system from subsurface conditions. Frash et al. (2017,2019) used simultaneous radiography/tomography measurements to compare mechanical and hydraulic apertures and found large disparities between mechanical and hydraulic apertures due to relatively closed segments of the fracture system (Figure 3).
Mature fault systems commonly contain both mature and immature fractures (Faulkner et al., 2010). These are characterized by a gouge-filled fault core that is relatively impermeable, which is bounded by mature slip surfaces. A surrounding damage zone contains relatively immature, low-displacement fractures and commonly possesses relatively higher permeability (Evans et al., 1997). Transport processes in mature fault systems are likely to be dominated by the complex architecture of the off-fault damage zone.

3.2.2. Coupled processes

Many experiments study the impact of coupled T-H-M-C processes on fracture permeability. In this section, we describe four such processes where much of the research has focused: how changes in normal stress (H-M), fracture displacement (H-M), and chemical reactions (T-H-C) affect fracture permeability and how solute/colloid transport (T-H-C) is impacted by fracture geometry. There are a number of other important coupled processes that have received less attention but which we lack space to discuss: the impact of fracture healing (T-H-time; Im et al., 2018; Song et al., 2021); fracture slip velocity (H-M; Jia et al., 2019); and particle mobilization (H-M-C; Candela et al., 2014).
Impact of normal stress on fracture flow (H-M)

The impact of normal stress is important to understanding how permeability changes for fractures in different orientations with respect to the stress field and how fracture permeability changes as a function of depth and fluid pressure. Fracture permeability decreases as effective normal stress increases due to strong hydromechanical coupling, which has important implications for managing production in unconventional oil and gas as well as geothermal energy. Experimental studies show that the application of normal stress results in an exponential decay of the mechanical aperture (Bandis et al., 1983), and a similar exponential decay of hydraulic aperture (Evans et al., 1997; Li et al., 2021; Skurtveit et al., 2020; Witherspoon et al., 1980). In addition, experimental studies of the permeability of shear fractures as a function of the normal stress during fracture creation show an exponential decay in permeability with increasing stress (Frash et al., 2017).

Normal stress leads to deformation of asperities, void morphology, and the adjacent matrix while altering flow path connectivity (Hopkins 1990; Pyrak-Nolte 2019). Goodman et al. (1968) developed an effective parameter, fracture specific stiffness, to represent changes in geometry. Research has shown numerically that fracture specific stiffness captures the deformed morphology of a fracture, and used these results to develop a nearly universal scaling relationship between fluid flow through a fracture and fracture specific stiffness (Petrovitch et al., 2013; Pyrak-Nolte 2019; Pyrak-Nolte et al., 2016). A challenge remains to demonstrate this relationship experimentally and difficulties arise from (1) the scaling relationship represents the ensemble behavior of a Monte Carlo study with 1000s of simulations, meaning that any single measured value of flow and stiffness in the laboratory may deviate from the averaged ensemble behavior; (2) the loading conditions on a sample need to be fully defined and (3) the critical neck, i.e. the smallest aperture along the dominant flow path, must be identified. Petrovitch et al (2014) showed that the loading conditions on a fracture affect the ability to develop a scaling relationship, especially for laboratory conditions. In experiments, care must be taken to avoid edge effects where the sample preferentially closes (c.f., a similar issue with permeability measurements; Petrovitch et al., 2014). To identify the critical neck, recent work combining X-ray imaging and magnetic resonance imaging (MRI) has the potential to extract the information.
needed to experimentally test this relationship for fractures under normal loading (Karlson et al., 2021). Additional research is also needed to determine the role of shear stress on shear fracture stiffness. Few studies have measured shear fracture stiffness and usually with ultrasonic methods (Choi et al., 2014). Recent x-ray radiography/tomography imaging capabilities developed by Frash et al. (2017, 2019) allows fracture morphology and permeability to be obtained for fractures that deformed under shear stress. The experimental results show that normal stress has a first-order effect on fracture permeability. Moreover, because we have good knowledge of normal stress in the subsurface, these results can be readily applied to applications requiring management of fracture permeability.

**Impact of fracture displacement on fracture flow (H-M)**

Barton et al. (1995) in a borehole study found that fractures oriented optimally with respect to shear stress (interpreted as more likely to experience recent shear displacement) had higher apparent permeabilities than non-optimally oriented fractures. This is attributed to the fact that fractures can heal over time and renewed displacement can rejuvenate permeability. Similarly, injection of fluids for waste disposal or hydraulic stimulation can reactivate fractures, potentially increasing permeability. Fracture reactivation and displacement can result in dilation that creates porosity as surface ride over rigid asperities but also creates fracture gouge as asperities are ground down.

Experimental studies of shear displacement of fractures are often conducted in conventional (Lee and Cho 2002) or triaxial direct shear devices (Crandall et al., 2017; Frash et al., 2016, 2017, 2019; Fang et al., 2017) as well as rotary shear systems that allow much greater displacement (Tanikawa et al., 2015). Changes in fracture permeability following displacement depend on the normal stress, rock type, and amount of displacement. At lower stresses with more brittle rocks, shear displacement results in dilation and fracture permeability generally increases (Lee and Cho 2002; Crandall et al., 2017; Frash et al., 2017; Ishibashi et al., 2018). As displacement increases, gouge accumulates eventually reducing or ending permeability increases (Lee and Cho 2002). As normal stress increases, dilation decreases such that permeability increases are small or negligible (Frash et al., 2017; Wang et al., 2021), while increasing phyllosilicate content (e.g.,
clays) results in decreasing permeability with displacement (Fang et al., 2018). The closely related phenomena of fluid pulsing (which changes the effective normal stress) results in enhanced permeability (Elkhoury et al., 2011). Further insights into fracture reactivation will require understanding the impacts of healing time (e.g., Im et al., 2018) on changes in fracture permeability as well as analyzing how displacement impacts fracture geometry (e.g, Zhao et al.’s 2018 rotary shear experiments).

Chemical reactions and fracture flow (T-H-C)

Field observations of fractures infilled with quartz or calcite are common and evidence of complex, fracture-hosted reactions are key features of many ore deposits. Dissolution reactions can enhance permeability by creating porosity or can reduce permeability by dissolving stress-weakened asperities; Precipitation reactions can fill porosity with precipitates, in reactions that are critical for the performance of geothermal systems and nuclear waste repositories. Chemical reactions will also affect fracture stiffness because the alterations may erode apertures or increase contact area and will deform under normal stress (Spokas et al., 2018). The location and relative geochemical reaction rates in natural fractures will affect how a fracture aperture is eroded, where heterogeneous precipitates will form and how the fracture deforms under stress. More research is needed on the impact of mineral distribution on fracture geometry in the presence of coupled processes. Fracture reactivity is studied in coreflood experiments and through the use of transparent analogs and microfluidics systems.

The kinetics of chemical reactions are such that dissolution reactions are more easily studied than precipitation. Polak et al. (2003) studied the friction and permeability response of a natural fracture in microcrystalline silica following injection of HCl acid. They observed substantial reductions of permeability due to dissolution of asperities confirmed by x-ray CT. On the other hand, experiments with HCl on highly reactive limestone showed in one case strong increases in permeability (Noiriel et al., 2013) and, in another case complex, wormhole evolution with strongly enhanced permeability with increasing flow rates (Noiriel and Deng 2018). In experiments that involve the potential for both dissolution and precipitation, Davila et al. (2016) found that in competing calcite dissolution and gypsum precipitation resulted in little change to
permeability, while Menefee et al. (2020) found extensive precipitation during competing calcite dissolution and witherite (BaCO$_3$) precipitation that concentrated in tight regions of the fracture and within comminuted fracture material (Figure 4). Menefee et al., observed relatively small to negligible decreases in permeability.

Figure 4: Scanning electron microscopy image of barium carbonate in a fracture network obtained from an induced shear experiment (shear fracture runs vertically through the specimen) involving dissolution of a carbonate-rich shale and precipitation of BaCO$_3$ (bright white) in response to coupled injection of BaCl$_2$ solution (after Menefee et al., 2020).

Transparent analog experiments of Detwiler (2008) showed a similar transition to channelized flow at higher flow rates and also showed dissolution of asperities and closing of the fracture. Microfluidics experiments of Jiménez-Martínez et al. (2020) examined dissolution and precipitation in supercritical-CO$_2$-exposed limestone and showed that preferential flow of the CO$_2$-phase through larger fractures homogenized dissolution reactions. We are just beginning to understand the complex dynamics of dissolution and precipitation. Fracture behavior is clearly strongly affected by the distribution of reactive mineralogy and interactions with fracture debris and stress-induced closure.
Chemical reactions also impact the matrix adjacent to the fracture and modify fracture stiffness (Figure 5; Eichhubl et al., 2004; Antonellini et al., 2017). The matrix material can be selectively weakened by the dissolution of specific minerals, resulting in irreversible compaction that can strongly affect fracture geometry. Other geochemical reactions may lead to strengthening of the reactive region from selective precipitation in matrix pores immediately adjacent to the fracture. Some studies have examined the extent of reaction halos but more data are needed to quantify the effect of halos on the hydromechanical behavior of fractures and elastic wave interaction (e.g., Figure 5).

Developing models of chemical reaction processes are difficult, particularly because of uncertainty of the effective reactive surface area. In fracture systems, this is compounded by creation of highly reactive fracture gouges. Kinetic barriers to nucleation and growth pose even greater obstacles to laboratory study. However, progress is being made, chiefly with dissolution, that allows an understanding of coupling between chemistry and flow that are critical to applications such as nuclear waste and geothermal energy.

*Figure 5:* Fracture (a) prior and (b) after seven days of flow of an acidic brine. (c) Irreversible compression of the reaction halo from the application of normal stress to the fracture plane. (with permission from K. Spokas 2021 - personal communication).
The previous discussion of coupled processes has focused on fracture permeability, but solute and colloid transport is likely to differ from simple single phase flow. For applications such as radioactive waste disposal, Rn and colloid-facilitated Rn transport must be assessed in performance assessment studies. Experimental studies of solute and colloid transport are quite limited, but transport is strongly affected by fracture geometry (e.g., stagnant zones), absorption/desorption processes, exchange with the adjacent matrix, and by non-flowing segments of the fracture network. Experimental measurements of solute production from fractures include those of Gale et al. (1990) and Stoll et al. (2019), where the latter attributed tailing of breakthrough curves to either roughness or aperture variations. A few studies have applied x-ray CT methods to visualize solute transport in fractures (Henry et al., 2014; Polak et al., 2003b,c). Fracture-matrix interactions have been investigated by Polak et al. (2003b,c) who found that fluid density created asymmetric diffusion profiles and that micro-fractures at the fracture-matrix interface enhance diffusion. The precipitation studies of Menefee et al. (2020) illustrate communication between the main fracture and the damaged matrix (Figure 4). The sorption behavior of tracers has also been used to interrogate fracture surface area (Dean et al., 2015; Roshan et al., 2016). Additional results on solute behavior have been obtained in computational studies (e.g., Cardenas et al., 2007; Kang et al., 2019).

Reviews of colloid (defined as particles from 10 nm to 10 µm) movement in saturated and unsaturated porous media (DeNovio et al., 2004; Ryan and Elimelech 1996) show that the factors affecting colloid deposition and mobilization include physiochemical attachment to grains, pore straining, attachment to air-water interfaces, entrainment in thinning water films, dispersion by chemical perturbations, expansion of water films, air-water interface scouring and shear mobilization. Of specific concern is how rapidly and how far particulates may be transported through the fractured porous subsurface. Fast transport of particulates can occur if they form swarms (Boomsma and Pyrak-Nolte 2015). Particle swarms are dilute suspensions of nano- to micro-scale particles that exhibit coherent behavior enabling a group of particles to travel 10 to 1000 times faster under gravity than single particles (Boomsma 2014; Boomsma and Pyrak-Nolte 2015; Machu et al., 2001a,b; Nitsche and Batchelor 1997). Such swarms often contain millions of colloidal size particles in a 5 µl drop. Because swarms result in enhanced speeds and
transport distances, contaminant swarms could result in higher concentrations farther from the source than if contaminants were released as a dispersion. In addition, particle swarms have been observed to reconfigure shapes to enable transport through narrow pore constrictions and tend to follow the dominant flow path (Molnar et al., 2019).

These experiments show that migration of solutes and colloids, which are important for understanding the fate of contaminants, requires additional analysis of fracture geometry and fracture-matrix interactions beyond determining flow lines to develop predictive models of transport in fractures.

3.3. New frontiers and future work

3.3.1. How lab measurements can and cannot inform field observations

For coupled processes in fractured rock, the purpose of laboratory experiments is to identify the fundamental physics and chemistry that affect fluid flow and transport through fractures through the use of well-designed experiments with controlled boundary conditions and initial conditions. This review demonstrates the rich complexity of even a single simple fracture and yet field studies (Section 2) and large-scale modeling studies (Section 4) involve observations and predictions of fracture networks. Nonetheless, laboratory experiments must provide a foundational understanding of coupled processes for single fractures, parallel sets of fracture and very simple fracture networks. However, to bridge to the field scale requires additional complexity and a range of length and time scales that cannot be achieved in the laboratory alone. The first step to bridging the gap between the controlled simple laboratory fracture systems and the massively complex field-scale fracture systems is through laboratory-validated computational simulations. It is only through simulations as described in section 4 that the complexity of the fracture systems can be ramped up and studied based on the fundamental physics learned from the data-driven experiments.

3.3.2. Laboratory experiments on fracture networks
Few laboratory experiments have been performed on fluid flow through fracture sets and fracture networks that are critical to interpreting field-scale phenomena. We do not yet have experimental data that demonstrates the relative importance of single fracture properties versus the connectivity controls exerted by fracture networks. On the other hand, computational studies (Section 4) and uncertainty quantification (Section 5) focus on the networks and typically neglect single fracture effects. These issues need to be resolved to effectively integrate experimental data with network modeling.

One of the key uncertainties in network behavior is the properties of fracture intersections (and how these are impacted by couple processes). Do these act as constrictions or openings relative to individual fractures? Fracture intersections also break the universal scaling in stress-permeability relations observed in two-dimensional single fracture behavior (Pyrak-Nolte and Nolte et al., 2016), and these intersections act as sources of seismic wave scattering that is not observed for single fractures or even for sets of parallel fractures. Fracture intersections are quasi-linear features with voids and contact area (Figure 3). Although the total void volume of intersections is only a fraction of the total fracture network volume, intersections control the connectivity of the network which can result in a transition from a fully 3D flow path geometry to lower dimensions (ultimately linear) with increasing stress. How this transition occurs depends on the void geometry of all of the basic elements that compose a fracture network and how they are connected. Field-tests in western France found that microbial hotspots at fracture intersections in the subsurface could be attributed to increased vertical connectivity (with depth) created by the intersections (Bochet et al., 2020). Recent numerical simulation studies on 2D fracture networks (based on outcrop data from the Bristol Basin in the UK) subjected to stress (Kang et al., 2019) showed that both stress magnitude and orientation strongly affects the aperture and resultant flow, and that stress can lead to preferential flow paths and anomalous early arrival of tracers. These intriguing results need to be confirmed for fully 3D fracture network systems through laboratory experiments.

Very few studies have examined fluid partitioning and streamlines at intersections in 3D. Most notable in these studies are the laboratory experiments on transparent “X” class topology for synthetic parallel-plate plate fractures (Hull and Koslow, 1986) and on replica casts (Johnson and
Brown, 2001; Johnson et al., 2006) of fractures in siltstone and in granite to study solute mixing at intersections. Through direct observation and numerical simulation, the research showed that fluid mixing in intersecting fractures is controlled by the spatial correlations of the aperture distributions in fractures. In simulations Johnson and Brown, (2001) and Johnson et al. (2006) observed that local pressure gradients near fracture intersections resulted in fluid flow opposite to the direction governed by the global pressure gradient. These local variations in pressure gradients arise from the statistics of the fracture aperture distribution and affect solute dispersion (transverse spreading of solute) and solute dilution (concentration of solute).

This raises several questions that could be addressed in the laboratory such as: (1) for monitoring, is the seismic response of a fracture network simply the ensemble response of all of the individual fractures within the network or will guided modes or other converted modes obscure the interpretation of fracture properties? (2) What are the mineralogic and geomechanical controls on formation, topology and geometry of intersections and intersecting fractures especially in the presence of pre-existing fractures (e.g. Figure 6)? (3) How do fracture intersections change under normal stress, shear stress and displacement conditions? (4) Can the properties of intersections among fractures be obtained in the laboratory or in the field? (5) Is global network topology sufficient to predict the physical and chemical response of a fracture network or must intersections be taken into account? (6) Is there an equivalent flow-stiffness relationship for fracture networks such that we could produce a three-dimensional map of fracture specific stiffness for a fracture network that indicates dominant flow paths through that network? There is a wealth of untapped information about fracture intersections that could be obtained from additional laboratory experiments.
Figure 6: (a&b) A 3D X-ray reconstruction of 12.5 mm x 13 mm region of a gypsum sample showing induced vertical fractures that intersect a pre-existing horizontal fracture. Reddish-brown are the void spaces in the pre-existing fracture. (a) Top of the vertical fractures and (b) bottom of the vertical fractures. From the analysis (table) the induced fracture preferentially formed over the voids within the pre-existing fracture. *(adapt with permission from El-Fil, 2021).*

3.3.3. Opportunities for model validation

A top priority is model validation of the hypothesis that fracture flow and transport can be determined from fracture geometry measured at the appropriate resolution. This will require
innovation in experimental methods to fully characterize fractures and reduce boundary artifacts, or the boundary artifacts will have to be incorporated in the simulations during model validation. From there, simulations that incorporate rock mineralogical properties as a basis for calculating the impact of coupled processes on permeability can be validated against experiments with the impact of normal stress being an obvious first target. While numerical models have significantly advanced the ability to include the geometric and topological complexity of fracture networks in flow simulations (e.g., Hu and Rutqvist 2020; Hyman et al., 2015), experimental data are needed on fracture intersections and flow path connectivity in response to stress and chemical processes in order to allow model validation. Breakthroughs in 3D X-ray tomography under ambient and subsurface conditions along with the advent of 3D printing that allows study of fractured networks with known, reproducible geometries provides a path forward for calibration of numerical models that simulate coupled processes in fracture networks. For example, Figure 7 shows a comparison of tomography and a finite/discrete element model to predict fracture network structure. The experiments were critical for validating the model and the model was critical to determining which parameters control the fracture patterns. Additional integrated experiment-model studies are crucial to understanding fundamental controls of fracture flow.
3.3.4. Application of machine learning

Machine learning (ML) is capable of extracting information from noisy signals and images, and can also mimic complex processes such as complex simulations. For this reason, ML has the potential to improve characterization, prediction of flow, and monitoring of changes in fracture properties induced by coupled processes. ML-enabled, automated and reproducible segmentation of x-ray CT fracture could improve the fidelity of fracture geometry. The simulation of flow fields based on fracture geometry could be aided by ML methods developed from the full Navier-Stokes solution. ML is already being developed to monitor the failure state of shear fractures (Johnson et al., 2021). Application of ML to laboratory ultrasonic transmission/reflection data and acoustic emission signals could tease out subtle changes in wave velocities or arrival times of different signal components to infer changes in fracture geometry.

Figure 7: Comparison of an X-ray image of a fracture network induced in direct shear experiment in shale (~100 mm in size) and a simulation of the failure process using the finite element discrete element approach (Hyman et al., 2016). Colors from red to blue indicate increasing values of strain.
For example, coda-wave analysis with active sources enables identification of subtle shifts in phases that indicate geometry changes because the trigger time, frequency content and source amplitude are known and controllable (Sang et al., 2020). Research is needed to determine if ML-enabled coda-wave analysis can be used for monitoring fracture networks under stress to infer opening/closing of propagations of the network. In addition, laboratory experiments are needed that work with uncontrolled sources that more closely mimic the unknowns with natural sources (earthquakes) and induced seismicity (e.g. hydraulic fracturing). These natural and induced sources can vary in amplitude, phase and frequency content that might make coda-wave analysis of a fracture system difficult. New chattering dust methods (Pyrak-Nolte et al., 2020) is a potential source for the laboratory scale that can generate 100s-1000s of signals for use in ML studies. Developing a ML method that can illuminate and/or monitor pre-existing fractures or other induced fractures from induced seismicity would take us a step closer to real-time monitoring of subsurface engineering activities.

4. Models of T-H-M-C coupled processes in fractured rock

The field tests and laboratory experiments discussed in the previous sections provide constraints on physics-based numerical simulations used to understand, predict and engineer optimal behavior in fractured systems. Decision-makers rely heavily on simulations, including in situations where results are expected in real time. There has been a longstanding debate on the fidelity required for numerical simulations of fractured rock and the answer very much depends on the fractured system being modeled and the quantities of interest. Numerical simulations of detailed laboratory experiments discussed in the previous section are constrained by mappings of the fracture or fracture network geometry and with measurements of coupled T-H-M-C processes. Recent field studies such as EGS Collab (Dobson et al., 2017), HFTS (Ciezobka et al., 2018), MSEEL (Carr et al., 2017), some of which are described in the field observation section, have a great deal of information that can be used to constrain high-fidelity simulations. However, for many other fractured subsurface systems, much less information about the underlying fracture network is available. Regardless of the particular site, however, the range of relevant spatiotemporal scales modeling coupled T-H-M-C processes have been notably prohibitive to construct efficient and effective computational models in the past (Bonnet et al., 2001; National
Research Council 1996; National Academies 2020). The choice of how to represent the particular features of the fractured medium, specifically the fractures, the network they form, and the background medium, leads to different classes of modeling approaches. The choice of modeling approach is also constrained by the resolution with which a site has been characterized. While previously limited in their scale and capabilities due to computational resources, the rapid increase in computational power over the last 30 years has allowed these simulations to drastically expand the resolution of spatiotemporal scales and coupled processes. These new simulations utilize the latest in high performance computing resources, meshing algorithms, and numerical schemes to represent the underlying fracture network and the relevant coupled processes. In this section, we provide an overview of the computational methodologies for modeling coupled T-H-M-C processes in fractured media that are currently in use and discuss recent developments.

4.1. Modern computational methodologies: recent developments and state of the art

In general, T-H-M-C coupled process simulations require consideration of both the fractures and the rock matrix. Modern computational methodologies for simulating coupled T-H-M-C processes in fractured rock can be grouped into two high-level categories (continuum and discrete approaches) based on the (1) representation of fractures, (2) interaction between fractures and the surrounding porous media, and (3) complexity of fracture network structures. The most common continuum approaches are the equivalent porous medium (EPM), stochastic continuum (SC), and Dual Continuum (DCM). The most common discrete approaches are discrete fracture network (DFN), discrete fracture matrix (DFM), channel/pipe-network models (CN). Figure 8 (adapted from Berre et al., 2018 and used with permission) provides a high-level schematic of these methods and how they relate to one another. The abscissa axis portrays the detail of the background medium representation and the ordinate axis portrays the detail of the fracture network representation. In the upper right corner is the true fractured porous medium that portrays the exact structure of the fracture network along with the heterogeneity of the porous medium, which can never be fully characterized at all length scales. The figure presents the trade-offs that are made when selecting a conceptual/numerical model of the system. Detailed fracture representations are usually relevant for predicting quantities like breakthrough times of
solutes whereas including a detailed representation of the background medium are useful in capturing mass transfer effects between fracture and matrix when processes like matrix diffusion dominate long term transport times. Another item to consider is the computational resources that are available; continuum models are typically less computationally demanding than discrete models. In addition to continuum and discrete models, there are also hybrid models that partition the domain into regions where an explicit fracture representation is used and others where a continuum method is applied (Dershowitz 2006; Joyce et al., 2014). While this allows detailed resolution of fractures close to an object of interest, e.g., a repository or well, and reduction in computational cost in the far field, the method is complicated by the need to communicate flow and transport properties between the subdomains.

Figure 8: Adapted from Berre et al., 2019 and used with permission. Schematic of trade-offs in fidelity of the fracture matrix versus the porous medium background in the representation of
a fractured porous medium for various numerical simulation approaches. Equivalent porous medium (EPM), stochastic continuum (SC), Dual Continuum (DC), discrete fracture network (DFN), discrete fracture matrix (DFM), channel/pipe-network models (CN) and hybrid discrete fracture network / equivalent porous medium (H-DFN/EPM).

It is also worth pointing out that all of these models are constructed stochastically to account for uncertainty in both fracture and matrix properties and extent, and therefore multiple realizations must be considered to address uncertainty and variability. The section following this one provides a detailed discussion of how these uncertainties are quantified, and we do not discuss it here. In the following subsections, we provide a brief discussion of the methods. Comprehensive reviews and comparisons between different approaches in the context of single-phase flow and transport in fracture networks are found in Berkowitz 2002, Berre et al., 2018, Hartley and Joyce 2013, Selroos et al., 2002, Ijiri et al., 2009, and National Academies 2020.

In addition to the network/field-scale models that we overview, there have been significant advancements in the direct numerical simulation of flow and transport at the single fracture scale which are based in the laboratory experiments discussed in a previous section (Bouffet et al., 2006; Cardenas et al., 2007; Detwiler et al., 2000; Kang et al., 2016; Lee and Kang 2020; Méhuest and Schmittpbuhl 2000; Lee et al., 2014; Lee et al., 2015; Zimmerman et al., 2004; Zou et al., 2015; Zou et al., 2017; Nicholl et al., 1999; Detwiler et al., 2007; Petrovitch et al., 2013). These simulations have demonstrated the relative impact of fracture surface roughness and fracture geometry against flow properties, such as Reynolds number, on the formation of in-plane flow channels, scaling relationships among geometry, flow and fracture stiffness, along with other small-scale properties and explored the deviations from the classical cubic law. A promising avenue of research on the horizon is bridging single fracture-scale observations from the previous section with those at the network scale and characterizing the hierarchy of scales across an entire network. In due course, there might be a need to integrate such small-scale effects into the larger-scale models that we discuss here as sub-grid scale processes.

Figure 9 provides images of modern 3D computational domains for flow and transport in fracture networks. The channel/pipe network, equivalent continuum model, discrete fracture matrix, and discrete fracture network models are all based on the same network and are generated using the
dfnWorks (Hyman et al., 2015) software. The photograph in the center (Precambrian Ortega Quartzite in New Mexico USA) highlights the complexity of the fracture network that modelers are trying to represent using the methods discussed in the next section.

Continuum Approaches (EPM, SC, and DCM):

In continuum methods fractures are not explicitly represented, but their effects on flow and transport are accounted for using effective or upscaled flow, transport, thermal and geomechanical properties. These methods are an obvious choice in situations where computational efficiency is key, quantities of interest are averaged over longer time scales, fracture network topology is poorly understood or a combination of these. In EPM approaches, fracture and matrix properties and fracture network structure are “homogenized” or “upscaled” to define gridblock-scale effective properties that describe the behavior of the entire fractured rock mass. Examples of EPM methods and applications are found in Botros et al., 2008, Hartley and Joyce 2013, Jackson et al., 2000, Reeves et al., 2008, and Sweeney et al., 2020. The distinguishing element of these models is how equivalent properties of the continuum are determined from the fracture and network attributes. Until recently, most upscaling methods were analytical, see (Liu et al., 2016) for a review of analytical expressions for upscaling porosity and permeability of fracture networks including the seminal work of Oda 1985 which is likely the most widely used analytical technique. However, simulation-based methods, where representative sub-grid scale DFN simulations can provide upscaled relationships including elements of the permeability tensor are becoming more popular (Bigi et al., 2013; Bogdanov et al., 200; Chen et al., 201; Durlofsky 1991; Karimi-Fard and Durlofsky 2016; Lang et al., 2014; Romano et al., 2020). These numerical methods will typically generate an ensemble of discrete models based on a site characterization at the grid block scale, e.g., a small DFN, then perform a set of flow simulations to extract parameters for use in the EPM. In essence, these numerical upscaling methods are a bridge between continuum and discrete models.
Figure 9: Examples of modern 3D models. All four models are based on the same fracture network. The photograph in the center of the Precambrian Ortega Quartzite in New Mexico USA highlights the complexity of the fracture network. Photo Credit: Jeffrey D. Hyman

SC and related heterogeneous continuum approaches allow for stochastic spatially variable representations of these effective properties, which may be informed by stochastic properties of the underlying fracture network (see Neuman 2005 for a comprehensive discussion along with references for examples therein). Hydrologic properties of an SC are represented using correlated random fields, and typically follow commonly used distributions e.g., Gaussian, Log-Normal, or multi-Gaussian, with correlation lengths derived from field observations. The SC method is particularly useful when information about specific fracture locations is scarce and a single realization of the random field will not represent a particular field site well, but can provide information about the expected ranges of bounding behavior. Most EPM and DFN/DFM models
also include elements of SC to account for variability and facilitate quantifying uncertainty, e.g., permeability in the continuum portion of the DFM can be represented by an appropriate random field via an SC model. Dual continuum models (DCM) also do not explicitly resolve fractures, but recognize the existence of a connected fracture continuum that overlaps and interacts with the rock matrix continuum. In the dual continuum model (dual-porosity/permeability), there exist two separate domains. Each domain is assigned unique properties. The fluxes of water, energy and solutes between the two continua are typically represented using linear exchange formulations and driven by pressure, temperature or concentration differences between them. Discussion and examples of DCM are found in Lichtner 2000, Moinfar et al., 2013, and Zimmerman et al., 1993.

Continuum models offer several advantages that are listed here: (1) the inclusion of both fracture and matrix properties, which is critical in many problems, (2) because fractures are not explicitly resolved, there are fewer computational and meshing constraints that limit computational domain sizes, and (3) the ease of use with conventional porous media simulators. As a result, the computational burden of continuum models is typically much smaller than discrete methods as fewer length-scales are explicitly resolved in the simulation. Additionally, most numerical schemes for coupled T-H-M-C processes, which are discussed later in this section, were developed in the context of the continuum framework. However, these processes can be extended to discrete methodologies as well. Finally, there are a wider range of continuum simulators available on the market (both commercial and open-source software) than discrete models. These aspects have favored continuum models to remain the standard methodology used today despite their limitations, which we discuss below.

Although continuum models appear simpler than discrete models, they present similar challenges in terms of theoretical foundations, numerical implementations, and uncertainty quantification. From a theoretical point of view, the definition of equivalent flow, geomechanical and transport properties in continuum approaches, often requiring the existence of a representative elementary volume (REV) which itself is an unclear notion (Neuman 2005) and can be difficult to define and identify. The process of upscaling requires the identification of a length scale that distinguishes explicit fractures from the background matrix and the resulting upscaled properties can be quite sensitive to this choice (National Academies 2020). For some fracture network topologies, the
existence of homogenized equivalent or effective properties at all is questionable (Bour and Davy 1997; de Dreuzy et al., 2001; Darcel et al., 2003; Molz et al., 2004; Neuman 2005; Sherman et al., 2019). There are some studies suggesting that these properties are scale dependent further complicating the identification of effective/equivalent properties (Neuman and Di Federico 2003). Similarly, the parameters used to construct the underlying fracture network used for upscaling in EPM or the selections made for the properties of the random fields in SC are subject to the same uncertainty concerns as when constructing a DFN/DFM model. Additionally, the analytic methods for upscaling typically do not consider the fracture connectivity within each grid cell, which is a critical aspect in fractured systems.

Depending on the mesh resolution of the continuum model, aspects of fracture network connectivity and anisotropic upscaled behavior may not be preserved accurately (Sweeney et al., 2020). Most upscaling methods are better at retaining spatial heterogeneity and anisotropic permeability than discrete pathways and network connectivity, which are critical aspects controlling coupled processes in fractured media. For example, if the grid resolution is too large, then false connections between fractures can be formed within a grid cell. These false connections have been shown to predict drastically earlier breakthrough of tracer when compared to discrete models (Hagdu et al., 2018). As with EPM, the effective properties of the fracture continuum in DCM may also be based on sub-grid scale fracture networks or other conceptualizations, and the same challenges arise in defining these effective properties and exchange coefficients for transfer functions (Litchner 2000). Furthermore, from a numerical point of view, the selection of the grid resolution can have a drastic impact on the model results both in terms of effective parameters and numerical artifacts, e.g., diffusion and dispersion (Benson et al., 2017). Thus, similar to discrete models, there are a large number of user selected parameters and conceptualizations used in model construction that must be carefully considered.

Discrete Models (DFN, DFM and CN):

Although as stated previously continuum models are the standard, their limitations coupled with improved computational power point to discrete models as the future to accurately capture T-H-M-C coupled processes. In discrete models, coupled processes are directly simulated on explicit representations of the fractures and the network that they form. The most common discrete
models are discrete fracture network models (DFN), channel or pipe network (CN) models, and discrete fracture-matrix (DFM) models. DFN models explicitly resolve the geometry and properties of fractures as lines in two dimensions or planar polygons in three dimensions but do not directly consider matrix interaction. The explicit representation of fractures comes at a higher computational cost than continuum models because of the range of length scales that are present in the system, which can span from the aperture roughness in individual fractures to the domain size (Adler and Thovert 1999; Bonnet et al., 2001; National Research Council 1996; National Academies 2020). However, the surrounding matrix is neglected, which limits their application to problems with where the effects of the surrounding rock matrix on flow is negligible. While there is no explicit representation of the matrix in a DFN model, particle-based methods, such as time domain random walk (TDRW), can be used capture the effects of matrix/thermal diffusion on transport through the network (Delay and Bodin 2001, Gisladottir et al., 2016, Hyman et al., 2019c, Hyman et al., 2020, Painter et al., 2008, Roubinet et al., 2012). DFN models originated in the 1980s and 1990s (Long et al., 1982; Cacas et al., 1990; Dershowitz and Fidelibus 1999), but were typically restricted to either 2D systems (Berkowitz and Scher 1997; de Dreuzy et al., 2001a; de Dreuzy et al., 2001b; de Dreuzy et al., 2002; Frampton and Cvetkovic 2009; Leung and Zimmerman 2012) or small 3D domains due to their high computational cost (Bogdanov et al., 2003; Koudina et al., 1998; Mourzenko et al., 2004; Painter et al., 2002).

Recent advances in both numerical schemes and computational power have given rise to the advent of large (tens of thousands of fractures in kilometer sized domains) DFNs (Berrone et al., 2013; Berrone et al., 2017; Benedetto et al., 2014; Davy et al., 2010; Davy et al., 2013; de Dreuzy et al., 2012; Erhel et al., 2009; Lang et al., 2014; Hyman et al., 2014; Hyman et al., 2015; Maillot et al., 2016; Pichot et al., 2010; Pichot et al., 2012; Thomas et al., 2020a). Despite being available through commercial, Fracman (Golder 2007) and ConnectFlow (Jacobs 2021), government, dfnWorks (Hyman et al., 2015), and academic (DFN.lab - Le Goc 2019; PorePy - Keilegavlen et al., 2021) sources, DFN models remain less commonly used than continuum models even with the advantages they provide over continuum models, which are discussed below.

Many of the early DFN models can be considered channel/pipe-networks (CN) (Long et al., 1982; Cacas et al., 1990; Moreno and Neretnieks 1993; Tsang et al., 1988). CN models may be
viewed as a precursor to modern DFN models, in that they represent the fracture network as one-dimensional pipe-like elements. Similar to a DFN, the fracture geometry is explicitly represented as one-dimensional elements that connect to form a series of channels through which the flow passes through. This reduction in dimension drastically reduces the number of degrees of freedom in the solution matrix, and thereby increases the computational efficiency. There has been a recent re-interest in these CN models, which are now commonly referred to as graph-based emulators (see section 5), as they allow for significantly more fractures to be included in networks compared to high-fidelity DFN models, which in turn facilitates and allows for many more simulations to be performed at low cost to bound uncertainty (Berrone, et al., 2020; Doolaeghe et al., 2020; Karra et al., 2018; Hyman et al., 2018; O’Malley et al., 2018; Osthus et al., 2020), a topic which is discussed further in the next section. The CN approach is attractive when there are a large number of fractures or the fracture intersections or solution enhanced permeability are a dominant factor in determining where flow and transport occur within the network. Moreover, CN methods readily allow for efficient particle-based modeling of matrix diffusion, sorption, and reactions via semi-analytical solutions.

Discrete fracture-matrix (DFM) models explicitly represent both the fracture network and surrounding porous matrix. While the general framework of DFM models dates back a few dozen years (Bogdanov et al., 2003; Koudina et al., 1998; Mourzenko et al., 2004; Painter et al., 2002), the development of DFM models for at scale application is relatively recent compared to the other computational models discussed here. This immaturity is primarily due to the complications associated with the conceptualization of the fracture network being represented using n-1 dimensional features embedded within a n-dimensional domain that represents the surrounding rock matrix. This formulation leads to several challenges associated with the computational geometry/mesh generation and numerical schemes for T-H-M-C processes. The computational meshes used in DFM models are inherently multi-dimensional, i.e., 3D volumes and 2D fracture-planes, and, in turn, novel numerical schemes must be developed for coupled processes to ensure enforcement of conservation laws. Berre et al., 2021 provides a series of benchmark exercises along with various numerical schemes in use today. Available DFM models can be loosely partitioned into two classes based on how they address the first of these two issues, which in turn directs how they address the second. DFM models either have a mesh
representation of the matrix that aligns/coincides with the mesh of the fracture network, referred to as a conforming mesh method (Ahmed et al., 2017; Bogdanov et al., 2003b; Helmisc 1997; Karimi-Fard 2004), or does not, referred to as a non-conforming mesh method (Berrone et al., 2018; Berrone et al., 2017; Boon et al., 2018; Flemisch et al., 2016; Frih et al., 2012; Fumagalli and Scotti 2013; Schwemck et al., 2015). To date, most DFM model studies are either 2D, quasi 3D, or used in relatively simple 3D geometries due to the relative infancy of the method and computational constraints.

Although computationally expensive, the main advantage of discrete models is that they provide a more faithful representation of gradients between the properties of fractures versus rock matrix (de Druezy et al., 2012; de Druezy et al., 2002; Edery et al., 2016; Frampton & Cvetkovic 2010; Hardebol et al., 2015; Hyman et al., 2019b; Hyman 2020; Kang et al., 2019; Maillot et al., 2016; Painter et al., 2002; Selroos et al., 2002; Zou and Cvetkovic 2020; Zou and Cvetkovic 2021) across a wide range of length scales (Bogdanov et al., 2007; de Druezy et al., 2002; de Druezy et al., 2012; Frampton & Cvetkovic 2010; Frampton et al., 2019; Hyman et al., 2019a; Joyce et al., 2014; Makedonska et al., 2016; Sweeney and Hyman 2019; Wellman et al., 2009) and offer the ability to test hypotheses about the importance of one scale of heterogeneity against another. This aspect makes them better suited to inform field site operators about what data would be the most impactful to constrain predictive models. On the other hand, even the latest and cutting-edge DFN/DFM simulators can only resolve tens of thousands of fractures, not the millions that are present at field sites, although it can be argued that fracture locations will rarely be known at a field site at all relevant length scales. DFM models are a reasonable compromise, allowing for the inclusion of large features, e.g., faults, to be explicitly included and the background fractures, whose distinction requires a user defined cut-off, to be accounted for using a continuum method. This hybrid formulation is a particularly promising avenue that we discuss later in this section.

4.2. Hydro-Mechanical (H-M) and Thermo-Hydro-Mechanical (T-H-M) Coupling

In this section, we review the application of these alternative modeling approaches in the specific context of T-H-M-C processes. In general, T-H-M-C coupled process simulations require consideration of both fractures and the rock matrix. The computational challenge of coupled process simulation has also restricted the ability to employ discrete fracture approaches,
especially for 3D problems. Nevertheless, advances in high-performance computation are putting these challenging problems within the reach of modern subsurface simulators.

In the absence of nonlinear dependence of material properties on stress, pressure and temperature, the standard porous continuum equations for linear thermo-poro-elasticity (Wang, 2000; Rutqvist et al., 2001; Cheng, 2016) provide a framework for modeling the evolution of pressure, stress and temperature fields in porous media. This set of equations includes the fluid flow and pressure diffusion equation, equations of mechanical equilibrium, and the energy transport equation incorporating advection and conduction. For fractured rock, T-H models in fractures are already quite mature so we do not present them here but can be found in (Wang, 2000). However, coupling mechanics (M) to T-H has seen many recent developments and are presented here. H-M and T-H-M coupled models need to represent the stress-dependence of fracture mechanical and hydraulic properties, and their evolution due to hydraulic or thermal transients. Although there is a more general class of problems involving fracture propagation and growth, requiring fracture mechanics formulations, we restrict this review to models that consider only pre-existing fractures and their evolution in response to stress and deformation. While H-M models consider the influence of stresses and fluid pressure on fracture deformation, T-H-M models also include thermal stresses. Changes in rock matrix temperature and associated deformation (contraction by cooling and expansion by heating) generate thermal stresses that augment hydro-mechanical stresses, leading to additional influences on fracture aperture, mechanical and hydraulic properties. Thus H-M models are a subset of T-H-M models. The importance of these coupled processes in fractured rock was highlighted by Tsang (1991) and Rutqvist and Tsang (2012). Previous reviews and surveys of H-M and T-H-M include Rutqvist and Stephansson (2003) and Lei et al. (2017).

The key ingredients for modeling H-M and T-H-M coupling in fractured rock are the nonlinear constitutive relationships for the evolution of single-fracture mechanical and hydraulic properties as a function of stress (or deformation) and pressure, which are summarized below.

Fracture Constitutive Relationships
In this section, we briefly present the equations for T-H-M-C processes in fractured media that are used by simulators. In general, the fracture aperture \( b \) at any location is defined as a function of the local stress tensor \( \sigma \) and the fluid pressure \( p \):

\[
b = f(\sigma, p)
\]

The normal \((\sigma_n)\) and shear \((\tau)\) stress components acting on the fracture plane are calculated from:

\[
\sigma_n = (\sigma_n).n, \tau^2 = |\sigma_n|^2 - \sigma_n^2
\]

where \(|.|\) is the magnitude of a vector.

Ideally, the initial fracture apertures must be calculated consistently based on the initial stress and pressure fields using the constitutive relationships shown below, which also describe the evolving stress-dependent and pressure-dependent aperture. The most commonly used constitutive relationship for normal deformation of fractures is the Bandis-Barton model (Bandis et al., 1983), which involves a hyperbolic dependence of fracture aperture on effective normal stress:

\[
b = b_{\text{max}} - \frac{A\sigma_n'}{1 + B\sigma_n'}
\]

where \( \sigma_n' = \sigma_n - p \) is the effective normal stress, \( b_{\text{max}} \) is the maximum aperture, the ratio \( A / B = b_{\text{max}} - b_{\text{min}} \), where \( b_{\text{min}} \) is the minimum aperture. The apparent normal stiffness of the fracture correspondingly increases with effective normal stress. Some variants of the Bandis-Barton model additionally distinguish mechanical versus hydraulic aperture through additional constitutive models (Min et al., 2003). Alternative models based on exponential rather than hyperbolic relationships have also been proposed (Min et al., 2003).

In general, enhancement of fracture permeability by shear dilation is greater than the increase due to reduction of normal effective stress (Min et al., 2003; Lei et al., 2017). There is much more variation among the approaches used to represent the influence of shear stress. Shear
deformation is typically modeled using elasto-perfectly plastic models, with elastic deformation up to a critical (failure) shear stress defined by the Mohr-Coulomb criterion (Min et al., 2004), at which point shear dilation is initiated and leads to large aperture growth. Most studies employ a critical shear stress \( \tau_c \) for shear failure based on the Mohr-Coulomb failure criterion

\[
\tau_c = \mu \sigma_n^f
\]

where \( \mu \) is the friction coefficient for the fracture plane, which is related to the surface roughness or joint roughness coefficient (JRC). The enhancement of fracture aperture by shear stress \( \Delta b_s \) is defined by the following relationships, which are based on the shear deformation \( u_s \):

\[
|\tau| \leq \tau_c: u_s = |\tau|/K_s, \Delta b_s = u_s \tan \phi_d1
\]

\[
|\tau| > \tau_c: u_s = |\tau|/K_s, \Delta b_s = (u_s \tan \phi_d2, \Delta b_{s,\text{max}})
\]

where \( K_s \) is the shear stiffness of the fracture, \( \Delta b_{s,\text{max}} \) is the upper bound on aperture enhancement, and \( \phi_d1 \) and \( \phi_d2 \) are dilation angles before and after shear failure. In many works, \( \phi_d1 \) is taken to be zero, i.e., shear dilation of fracture aperture is only assumed to occur after failure. However, stochastic analysis of shear displacement between self-affine random surfaces indicates the possibility of dilation in both regimes, which in general leads to anisotropic enhancement of fracture permeability (Mallikamas and Rajaram, 2005). Several other works (e.g. Rinaldi et al., 2015) employ a simplified version of the above relationships where there is an abrupt permeability enhancement at Mohr-Coulomb shear failure. In other words, the thresholded linear relationship in the post-failure stage is replaced with exponential saturation functions (Min et al., 2003; Cappa and Rutqvist, 2011).

Mineral precipitation and dissolution equations are used to modify the fracture aperture that can be related to permeability with equations such as the cubic law (see section 3.2.1). An example mineral reaction is given by:

\[
\frac{dc}{dt} = k(T)A\left[1 - \frac{Q}{K(T)}\right]
\]
where C is the concentration, k is the reaction rate constant, A is the surface area, Q is the equilibrium quotient and K the equilibrium constant. K and k are strong functions of temperature (Lichtner 2000).

4.3. Recent T-H-M-C Coupled Process Findings from Modeling Studies

Experimental verification of these constitutive relationships has primarily been conducted by laboratory experiments at the single fracture scale like those described in the laboratory experiment section. Recent modeling studies utilizing these types of constitutive relationships demonstrate sites such EGS Colab, fault slip studies in Mont Terri described in the field observation section qualitatively capture T-H-M-C feedbacks that have been observed in the field for decades. Acknowledging limitations in fully capturing T-H-M-C effects, these types of constitutive relationships are essential for modeling and at present they at least provide a phenomenological option that should be evaluated further.

Modeling T-H-M coupling in large fractured rock masses requires the integration of the aforementioned constitutive relationships with flow, temperature and geomechanical models for the entire rock mass. ECM and SC models attempt to represent the evolution of grid block-scale effective stiffness and permeability tensors and porosity based on assumed sub-grid scale fracture networks (Rutqvist et al., 2003). Min et al. (2004) presented numerical simulations of hydro-mechanical coupling in fracture networks to derive upscaled stress-dependent permeability tensors. These tensors represent the aggregated influence of normal and shear deformation of individual fractures described above, including the important contribution of shear slip along favorably oriented fractures, which contributes significantly to permeability anisotropy. In time-evolving T-H-M systems, grid-block scale effective properties would be expected to evolve with time. Few ECM studies have demonstrated the ability of stress/pressure-dependent effective properties to capture such evolving behavior and the emergence of preferential flow as a result of such evolving behavior (e.g., in the case of shear stimulation). The definition of evolving effective properties becomes problematic when evolving network structures have characteristic length scales much larger than the grid block scale. DFN and DFM
models do not suffer from this issue since they explicitly consider network structure. Additionally, the preservation of the true fracture network geometry and connectivity in a continuum model depends on the mesh resolution, and connectivity resulting from small-scale features that will be omitted if their characteristic length scale is smaller than the mesh cells that contain them as previously discussed. The strong contrast between fracture and matrix properties further challenges equivalent continuum models. For example, recent studies (Haagenson and Rajaram, 2021) suggest that although the propagation of spatially averaged fields (e.g., radially averaged pressure, temperature or concentration) may be accurately captured in equivalent continuum representations, propagation of high pressures and temperature perturbations, which drive stress-dependent fracture behavior and shear slip, which are the key elements of H-M coupling, could be greatly mis-represented by equivalent continuum models (Zareidarmiyan et al., 2021). Again, discrete methods can more accurately capture the propagation of these high pressure and temperature perturbations since network structure is explicitly considered, as clarified in the above studies.

An approach for DC modeling of T-H-M-C processes in a fractured rock mass using the TOUGHREACT and FLAC codes was presented by Taron et al. (2009) and Taron and Elsworth (2009). This approach was applied to a prototypical EGS system and successfully demonstrated the phenomenology of various coupled processes and the relative importance at different time scales. Specifically, these include permeability enhancement by short-term H-M mechanisms and intermediate-term cooling of the rock mass, and slow permeability increase/decrease due to mineral dissolution/precipitation in fractures. The DC approach incorporated single fracture constitutive relationships similar to those described above in the fracture continuum. However, no relationship to network structure and influence of fracture connectivity on effective properties of the DC system was incorporated. Since fracture structure is likely an important factor, including it is a next logical step. Such an extension of their modeling approach is feasible and would provide a framework for DC modeling of T-H-M-C processes in complex fracture rock masses, by incorporating information on the underlying fracture network structure. Lichtner and Karra, 2014 developed a new scalable algorithm for DC modeling thermal and reactive transport processes, and applied it to CO₂ sequestration and EGS applications. They showed that this framework allows for a general multi-component aqueous system with mineral
precipitation/dissolution. Furthermore, they showed that the additional computational overhead due to fracture plus matrix grid nodes is negligible compared to ECM.

The formulation of T-H-M-C coupled processes for simulating flow in fractured rock for discrete systems has been developed but has only been implemented for simple fractured systems. Discrete fracture models of T-H-M-C coupled processes in fractured rock require explicit representations of both the fracture and rock matrix. Two classes of discrete fracture modeling approaches have been developed for H-M and T-H-M problems – (a) Discontinuity formulations and (b) Discrete fracture network and matrix (DFM) models. Fully H-M coupled models in both classes compute stresses and pressures in the rock matrix using the equations of poroelasticity or thermo-poro-elasticity. Discontinuity formulations explicitly treat fractures as discontinuities and solve for the evolution of these discontinuities as a result of kinematic displacement and rotation of rock matrix blocks, which are treated as elastic continua. The Discrete or Distinct Element Methods (DEM; Min and Jing, 2003; Jing and Stephansson, 2007; Itasca Consulting Group, 2013) and Discontinuous Deformation or Displacement Methods (DDM; Shi and Goodman, 1985; Zhou and Ghassemi, 2011; Safari and Gassemi, 2015) are popular discontinuity-based formulations. In these approaches, displacements along fracture surfaces are used to calculate evolution of fracture properties.

In DFM models (Rutqvist et al., 2009; Figueiredo et al., 20015; Pandey et al., 2017), fractures are not treated as discontinuities but as embedded continua with an assigned numerical width that differs from the fracture aperture and mechanical and hydraulic properties that differ from that of the surrounding rock matrix. The mechanical and hydraulic properties of fracture elements or grid cells are assigned to consistently represent the behavior of fractures according to the aforementioned nonlinear constitutive relationships. In the most sophisticated DFM models, mechanical properties of fracture grid blocks are typically assigned to be weaker than that of the surrounding rock mass, while permeability along the fracture plane is orders of magnitude greater than that of the surrounding rock matrix (Rutqvist et al., 2009). In some DFM formulations, the stress calculations do not account for the presence of fractures, but the stress dependence of fracture aperture and permeability are incorporated in flow and transport calculations (Fu et al., 2016). Approximate uncoupled or partially coupled models do not
explicitly solve for the stress field (Birdsell et al., 2018; Haagenson and Rajaram, 2021), but only track the influence of pressure on effective stress and fracture properties (the stress field is externally specified, usually time-invariant). All of these approaches have been used to study problems ranging from the single fracture scale to small numbers of fractures, including 3D treatment of geomechanics in many works (Pandey et al., 2017; Salimzadeh et al, 2018). However, simulations on complex fracture networks with numerous fractures are still lacking. Admittedly, there are uncertainties introduced by the phenomenological representations of fracture constitutive relationships, and more comparisons with field observations will help to refine these representations further.

The role of chemical reactions such as precipitation and dissolution (the C in T-H-M-C) is to alter fracture aperture (and the matrix material immediately adjacent to the fracture see Figure 5)), which then influences flow and transport. Notably, fracture permeability is very sensitive to aperture, with a nearly cubic dependence assuming a local cubic law. Reactive alteration of fractures is a relatively slow process due to the large disparity in molar density of mineral components in solid versus aqueous phases (Ortoleva, 1987; Lichtner, 1988; Hanna and Rajaram, 1998). In dissolution-dominated regimes, unstable growth of dissolution fingers leads to increasingly preferential flow through fractures (Hanna and Rajaram, 1998; Szmyczak and Ladd, 2006) and self-organization. This self-organizing phenomenon has been invoked to explain the formation and morphology of meteori karst and cave systems (Howard and Groves, 1995; Hanna and Rajaram, 1998; Rajaram et al., 2009), and experimentally verified in analog systems (Detwiler et al.2003, Detwiler and Rajaram, 2007). Extension to turbulent flow regimes in fractures in the advanced stages of dissolution growth leads to additional morphological patterns resembling cave structures (Howard and Groves, 1995; Rajaram et al., 2009). Reaction rates and equilibrium constants are strong functions of temperature adding another critical feedback that needs to be considered for applications such as nuclear nonproliferation, geothermal and nuclear waste disposal where large temperature gradients exist. The reactive-permeability feedback is typically accounted for in a model as a change to the fracture aperture and mechanical properties of the matrix adjacent to the fracture as described in the constitutive relationships above. These reactions must often be treated kinetically since they are often slow and play an important role for how flow and transport at geothermal, carbon sequestration and
nuclear repository sites evolve over time. The reaction rate is strongly dependent on surface area and kinetic parameters which are often difficult to constrain. Nonetheless, researchers have made great strides in measuring kinetics and surface areas for many of the critical reactions (Steefel, 2018). The relatively slow rate of reactive fracture alteration facilitates lagged or sequential coupling of aperture alteration with T-H-M and advective-diffusive transport processes for computational efficiency and is thus readily incorporated into models with T-H-M-C capabilities. Chaudhuri et al. (2008, 2009, 2013) employed this approach to demonstrate several feedbacks resulting from aperture alteration driven by T-H-C processes, including the anisotropy resulting from precipitation, onset of buoyant convection due to aperture growth and reorganization of buoyant convection cells. These studies supplemented the advanced T-H-C capabilities already available in the FEHM simulator (Zyvoloski, 2007), by implementing aperture alteration.

It is also important to include fracture-matrix interactions in T-H-M-C models, which can significantly influence the species concentrations that drive fracture alteration reactions and alter rock matrix mineralogy and affect the local strength of the matrix. Alterations of rock matrix mineralogy is relevant to long-term evolution of radionuclide sorption surrounding nuclear waste repositories. Steefel and Lichtner (1998 a, b) presented an analysis of multi-component reactive transport in a single fracture-matrix system, illustrating the formation of mineral alteration fronts in the rock matrix, with applications to hyperalkaline plume transport. Pandey and Rajaram (2016) presented multi-component reactive transport simulations of weathering in the critical zone using the PFLOTRAN simulator, to illustrate the diffusive propagation of reaction fronts into matrix blocks and its implications for system-scale reaction rates. Chaudhuri et al. (2013) present a general formulation for incorporating fracture-matrix interactions including advective fluid exchange in T-H-C models of reactive fracture alteration. In the DFM class of fractured rock models discussed above, these fluxes are automatically calculated on the overall grid and can be readily included in transport computations. In very low permeability rock, discretization of the rock matrix at a resolution fine enough to accurately capture concentration gradients at fracture-matrix interfaces is a significant challenge in extending these models to large 3D T-H-M-C problems. Advances in high-performance computing technologies can provide some relief here. However, in many instances, tractable computation will require simplified hybrid models, such as one-dimensional streamline based reactive.
Other examples where chemical processes play a large role in subsurface system performance include hyperalkaline plumes at nuclear waste repositories (Viswanathan et al., 1998; Steefel and Lichtner, 1998b), barite precipitation filling fractures reducing production efficiency from unconventional reservoirs (Xiong et al., 2020), salt-precipitation in fractures at EGS systems (Borgia et al., 2012), and self-sealing of fractures the limit wellbore leakage from carbon sequestration reservoirs (Guthrie et al, 2018), and long-term reactive alteration, resulting feedbacks and implications for energy production rates in model EGS systems in various mineralogies (Rawal and Ghassemi, 2014; Pandey et al., 2014, 2015). Steefel (2018) provides a thorough review of reactive transport models of subsurface energy and environmental problems with some of the applications being fractured systems. A discussion/comparison of currently available models for reactive transport is provided in Steefel et al., 2015.

4.4. New frontiers and future work

With the confluence of better field and laboratory measurement techniques and numerical models, the field is well-positioned to include T-H-M-C processes into discrete DFM models that improve the mechanistical representation of fracture network evolution and its impact on flow and transport. Despite advances in field and laboratory techniques, the subsurface remains opaque, heterogeneous, uncertain and difficult to characterize. For this reason, numerical models are often data limited and not limited by the numerical schemes that have been implemented. The next section discusses the powerful tool of uncertainty quantification that partially remedies this problem. However, recent advancements in simulation provide a framework to test hypotheses to determine the critical mechanisms that control subsurface systems. Forecasting and prediction still remains elusive but by combining the mechanistic models in this section with the uncertainty quantification methods in the next section moves the modeling community towards this goal.

Multi-scale simulation is a new frontier for fracture simulation. Modern simulators now allow for multiple length scales of heterogeneity to be represented within field-scale models, at least for fluid flow, and probe their relative importance. For example, aperture variability/surface roughness (Makedonska et al., 2016; de Druezy et al., 2012; Frampton et al., 2019), can now be
included into large discrete model simulations, thereby providing a bridge between the domains of interest as discussed previously in this section. These simulations demonstrated that the fracture network structure was the dominant control of flow and transport behavior in most cases. Thus, one takeaway message moving forward is that DFM models for T-H-M-C coupled processes may not need to explicitly resolve single-fracture scale aperture variability, especially when the influence of the rock matrix is included.

At the slightly larger scale of individual fractures, fracture orientations and effective hydraulic apertures are also now known to affect flow structure within networks and lead to flow channelization even in the absence of internal aperture variability (e.g. Xu, 2020). At an even larger scale network density and connectivity have been found to alter the distribution of flow within the network (Hyman 2020; Lang et al., 2014; Sherman et al., 2019). In the context of T-H-M-C coupled processes, the next step would be to incorporate evolving fracture properties into these simulations.

All these developments in flow and transport generate optimism for the development of DFM simulators for T-H-M-C coupled processes in complex fracture networks. Although detailed 3D H-M and T-H-M modeling on complex fracture networks is only at its inception, previous studies on simpler 2D and 3D systems enable extension to complex 3D fracture networks. Several modeling studies (Kohl et al., 1995; Ghassemi and Zhou, 2011; Zhou and Ghassemi, 2011; Guo et al., 2016; Pandey et al., 2017; Salimzadeh et al., 2018) have considered the coupled T-H-M behavior of an extensive single fracture embedded within a 3D rock mass in a model EGS system. These studies considered the influence of pressure and thermal stresses on fracture aperture and permeability using discontinuity and DFM models. These studies were able to reproduce the significant flow channeling resulting from thermal contraction of the cooling rock mass, and a stress-cage effect that creates compressive stresses outside the region of cooling. Fu et al. (2015) and Sun et al. (2021) present 2D T-H-M models of EGS systems based on a DFM approach, and demonstrate the influence of network-scale flow channeling on short-circuiting. Safari and Ghassemi (2015) presented a simple 3D displacement-discontinuity formulation for T-H-M coupling in fractured rock masses. In addition to short-circuiting, Safari and Ghassemi (2015) and Sun et al. (2021) also demonstrate the ability to model the occurrence of shear slip
and induced seismicity. These processes lead to reduction in the energy production rate from the EGS as widely documented in field observations, and is commonly referred to as short-circuiting (see Section 2 EGS Collab). These studies provide encouraging proof-of-concept that commonly employed constitutive models for fracture deformation due to geomechanical and thermal stresses can reproduce behavior observed in real-world EGS systems. Future study of more complex networks with one-way couplings between the stress field and fracture properties can be used to understand the interplay of stress with macroscale channeling (Sweeney and Hyman 2019, Kang et al., 2019).

Another promising avenue of development in discrete models was made possible by computational and theoretical advancements is so called “kinematic” DFNs (Davy et al., 2010; Davy et al., 2013) and extended by recent work of Thomas et al., 2020a, Thomas et al., 2020b. Traditional DFN models are generated stochastically from distributions determined by field site characterization. In contrast, these kinematic DFN mimic the growth of fractures using simplified rules of nucleation, growth, and arrest. While lower fidelity than a full geomechanics based simulation, which is infeasible for thousands of fractures, it creates fracture networks with more physically realistic structure than a purely stochastic network, e.g., T-shaped intersections and naturally emerging power-law distribution of fracture lengths. For example, if a smaller fracture grows to intersect a larger one, the smaller fracture stops propagating and its terminus is a T shaped intersection, which is common in nature but nearly impossible to create in a fully stochastic DFN. These changes in network generation structure have been shown to influence the topological (Lavoine et al., 2020; Hope et al., 2015) and hydrological properties of the networks (Maillot et al., 2016). In particular, kinematic networks show increased levels of flow channelization than their stochastic counterparts. While these kinematic networks are far cheaper to construct than performing a full geo-mechanical simulation of fracture nucleation and growth, they are still significantly more costly than creating a purely stochastic DFN. Thus, the development of an emulator for kinematically generated networks, similar to a pipe-network, capable of system-scale self-organization resulting from coupled processes would drastically facilitate improved uncertainty quantification, as discussed in the next section. As more computational power comes online, these kinematic DFN are poised to play a larger role in DFN
simulations. Moreover, these kinematic DFNs could be used in continuum models via the numerical upscaling techniques discussed above.

Including chemical processes into DFM models does present some challenges. Often simulators represent mechanics with finite element formulations but chemical processes use finite volume formulations. Therefore, dual-meshing has been used, where vertices of control volumes employed in finite volume discretizations for flow and transport are used as vertices for stress calculations (TOUGH-FLAC (Rutqvist 2011), PFLOTRAN (Lichtner et. al 2015), dfnWorks (Hyman et al., 2015), FRACMAN (Golder 2021). Fully finite-element approaches (MOOSE – Gaston 2009) employ the same mesh for all variables, with mixed finite element formulations for transport. Both of these approaches have been proven and are a promising path forward for 3D T-H-M-C processes.

As noted above, most DFM model studies to date are either 2D, quasi 2D, or used in relatively simple 3D geometries due to the relative infancy of the method and computational constraints. Modern high-performance computing resources can be leveraged to accomplish extension of DFNM approaches to simulation of T-H-M-C coupled processes. The key advantage that DFM models carry over the others in use today is that large-scale features, such as faults, can be explicitly included in the domain and the background fractures can be upscaled with matrix properties into effective properties. This methodology facilitates the explicit inclusion of dominant pathways, which keeps the computational cost low even when the domain of interest is multiple kilometers in length. In turn, DFM can provide an avenue for the simulation of coupled T-H-M-C processes in large, complicated 3D fractured rock masses. With newly available static and dynamics measurements from field experiments and controlled laboratory experiments, future models should be able to predict trends observed from field observations and have the potential to improve prediction of key quantifies of interest such as heat or hydrocarbon extraction rates, leakage rates from carbon, hydrogen and nuclear waste repositories. Field tests in Section 2 have demonstrated that T-H-M-C coupled processes and fracture-matrix interactions are important for many subsurface applications and these can be mechanistically represented by DFM models.
As is evident throughout this section, the available computational models for coupled processes in fractured media are becoming more advanced in terms of their numerical schemes, spatiotemporal resolution, and the coupled processes they can resolve. The inclusion of additional physical processes such as multiphase flow or particulate or chemical transport and reactions into the most advanced DFM models will likely require novel mixed-dimensional discretizations for their respective governing equations. This requirement is somewhat worrisome as legacy continuum codes for coupled processes will not be directly accessible. Therefore, attention needs to be given to the transfer of knowledge between legacy codes and the next generation of models to ensure that the information produced over the last few decades will be retained and utilized as we move forward in our modeling efforts. However, many of these numerical methods are far ahead of the site characterizations discussed in Section 2. Therefore, the primary issues associated with numerical modeling of T-H-M-C processes in fractured media for predictive applications are no longer mathematical challenges nor are they computational limitations, but instead is how to constrain the models based on the limited, albeit growing information from the field. Despite highly detailed measurements at dedicated field testing sites, most sites have much less characterization data. Therefore, integration of field measurements and observations with numerical simulations are needed to determine the minimum set of characterization data needed to predict critical quantities of interest such as plume evolution and breakthrough time for a solute. In addition, research is needed to explore the transferability of data from a well characterized site to less characterized sites to help fill gaps in understanding. Machine learning methods such as transfer learning have shown promise for tackling this important problem but might be challenging for the subsurface sites that vary geologically and tectonically. Regardless of these advancements in computational methods for coupled processes in fractured media, the uncertainty inherently surrounding the characterization of fractured media remains a dominant consideration in model selection and design. Addressing that uncertainty is the focus of the next section. Combining uncertainty quantification approaches with these mechanistic models is critical for moving the field forward.

5. **Uncertainty quantification, simplified models and emulators**
Rapid advances in computational power have provided an impetus for the development of more detailed models of flow and transport, and more broadly coupled thermo-hydro-mechanical-chemical (T-H-M-C) phenomenon in fractured rocks. The previous section described breakthroughs in high fidelity physics-based models that can represent a large number of fractures of multiple length scales explicitly while also representing the fracture-matrix interactions. Although the latest field and experimental techniques provide better constraints in these models on fracture location, density, orientation and coupled processes at well instrumented sites, such information is typically not available for many applications. The lack of sufficient information leads to large uncertainty in the simulated behavior of systems of interest because bounds on these fracture parameters are not known in sufficient detail to constrain a fracture model. Such models (section 4) rely on a large number of (spatially and temporally varying) parameters that can never be ascertained with sufficient certainty (Tartakovsky and Winter, 2008) for the vast majority of sites. As discussed in the previous section, continuum models smooth out the effects of fracture connectivity/topology/geometry which in many circumstances is a critical mechanism that controls flow and transport through a subsurface site. In addition, determining which fractures actually participate in the flow adds more uncertainty. Therefore, despite being difficult to constrain, the DFN and DFM models are necessary for accurately simulating the flow physics even when all of the fracture detail is not known in its entirety.

In reality, fracture characteristics can never be fully known but are instead represented in a statistical sense. Consequently, quantitative predictions derived from such statistical representations of fracture systems must be accompanied by confidence intervals, i.e., by a robust uncertainty quantification (UQ). The goal of UQ is to obtain probabilistic estimates on key quantities of interest such as travel time through the fractured system and the temporal evolution of the solute plume. With a few exceptions, such as stochastic finite elements (Elman, 2017) and the method of distributions (Tartakovsky and Gremaud, 2017), UQ methods involve repeated solutions of the governing equations (e.g. Monte Carlo) with different combinations of input parameters. The error in such ensemble-based simulations has two sources: the error due to the reliance on the finite number of forward model runs (aka realizations) and the model error caused by the inevitable approximations “reality”. Since higher-fidelity models are typically
more expensive, this error balance introduces a trade-off between a model’s fidelity and its computational cost: given a fixed simulation budget, an optimal model is not necessarily (if ever) the one that incorporates most physics (Sinsbeck and Tartakovsky, 2015; Yang et al., 2020) but is the one that fits into the budget. Ensemble-based computations are also employed in inverse modeling, most commonly used to estimate the parameters in models and in data assimilation to update the parameters in a model. However, most decision-making exercises seek to invest resources in the computation of uncertainties instead of ascertaining the “best” models/parameters because parameter estimation methods require significant computational resources and the “best” parameters still do not result in a high degree of confidence in the models due to the overwhelmingly heterogeneous nature of subsurface systems.

In many cases, the DFM and DFN models are already computationally very expensive for just one realization and thus running 1000s of realizations is not feasible at this point in time. Clearly there is a need for balance between fidelity and computational cost, and this balance depends on how well information from field and laboratory tests can constrain a particular site. Sites with little characterization and validation data may not warrant high-fidelity models but instead rely on simple continuum models with an emphasis on the calculation of robust uncertainty bounds. Continuum models are also adequate if the fracture density is sufficiently high in each grid cell such that it is better approximated by an effective continuum than trying to include a massive number of discrete fractures. Well-characterized sites with detailed fracture information can take advantage of high fidelity models but may still require many realizations to bound a system’s behavior.

A more desirable approach is to improve the fidelity of the physics in a model while simultaneously decreasing the computational burden (Figure 10). This can be achieved by coupling Machine Learning (ML) into simulation workflows which has the potential to be groundbreaking in its ability to improve agility and innovation. ML algorithms are increasingly being incorporated within the user workflow to accomplish two objectives: (i) to learn surrogate models to speed up computations by orders of magnitude and (ii) to find hidden patterns in high dimensional model outputs that are virtually impossible for the human brain to detect, even with vast subject domain expertise. ML methods have shown great promise in a variety of
applications such as genomics (Libbrecht and Noble, 2015), cancer research (Kourou et al., 2015), material science (Butler et al., 2018), drug discovery, social media (Golbeck, 2011), cybersecurity (Buczak, 2015), and more recently COVID-19 research (Zoabi et al., 2021; Randhawa et al., 2020). The data in these diverse fields share common characteristics with data from geoscience applications, namely high dimensionality, high spatio-temporal correlation and heterogeneity, multi-source multi-resolution data and sometimes small sample size, poor quality data, and occurrence of rare events. ML is very well suited for the characterization of flow and transport in fractured rock because high-fidelity DFNs can be both emulated through efficient surrogates and reduced through the detection of primary flow pathways and mechanisms. A common criticism of ML algorithms is that when forced into extrapolatory regimes, they often result in predictions that are fraught with uncertainty. To overcome this limitation, scientists have resorted to moving away from purely data-driven approaches to physics informed or physics constrained approaches wherein the algorithm incorporates the underlying physics of the system. When ML algorithms have been trained with ample data and properly constrained by the underlying physics, they have the ability to capture the physics more accurately and to achieve, simultaneously, orders of magnitude decrease in computational time.
In this section, we describe a combination of simplified stochastic models and emulators that enable robust uncertainty quantification for fractured systems. Simplified stochastic models and emulator approaches calculate the confidence intervals on model outputs in different ways. Simplified stochastic models make numerous simplifying assumptions that often result in analytical or semi-analytical solutions with a few critical lumped parameters that describe the behavior of a fractured system. An advantage of these models is that their elegance makes it easy to gain intuition about a fractured system. However, the results must be examined carefully because many assumptions are made that need to be evaluated to determine if they apply to the...
system of interest. Emulators take advantage of recent advances in physics-informed machine learning and graph theory to accelerate the physics-based models described in the previous section by 3-4 orders of magnitude which enables many realizations that contribute to a robust uncertainty quantification. The emulator approach has the advantage of being anchored by the high-fidelity physics-based models. Both stochastic models and emulator approaches show promise and are detailed in this section. So far, these simplified stochastic models and emulators have only been developed for single phase flow. A much needed next step is to Extend these reduced-order modeling approaches to more complex coupled T-H-M-C processes and to build UQ frameworks that encompass these models and evaluate the uncertainties associated with the various individual processes and their coupling in different subsurface applications (Figure 1). Along with the recent advances in UQ for flow and transport in fractured systems, we describe the challenges and possible paths forward for T-H-M-C related development.

5.1. Recent significant advances

5.1.1. Flow and transport in fractured rocks: simplified models for UQ

A hierarchy of conceptualizations of flow and transport in fractured rocks exists that is vast and varied in accuracy and complexity, and that continues to evolve (see, e.g., Berkowitz, 2002; Neuman, 2005; Berre et al., 2019; National Academies Report, 2020; among other reviews). The following examples serve to highlight the variety and various levels of mathematical abstraction, but an exhaustive description of such models is beyond the scope of this paper. Arguably the most simplified representation of a fractured rock is to treat it as a (stochastic) continuum (e.g., Neuman, 1988; Hyun et al., 2002; Botros, 2008), which does not distinguish between fractures and the ambient matrix. This is conceptually similar to Darcy’s model for a porous medium, but accounts for the potential absence of a representative elementary volume. Multicontinua and mobile-immobile models (Goltz and Roberts, 1986; Lichtner and Karra, 2014; Al-Rudaini et al., 2020) represent the next level of complexity to account for the presence of preferential flow paths in fractured rocks, while staying within the Darcian framework. Hybrid discrete-continuum models explicitly represent dominant fractures, while representing smaller fractures as a continuum (Carrera and Martinez-Landa, 2000; Fisher et al., 2020). Finally, discrete fracture network (DFN) models attempt to describe individual fractures and processes occurring in them.
in a manner akin to pore-scale modeling (Dverstorp and Andersson, 1989; Fumagalli et al., 2019).

It is worth noting that a conceptual model with a more “realistic” representation of fractured rocks does not necessarily have a higher fidelity. That is because more complex models invariably involve more unmeasurable parameters and are often underpinned by simplifying assumptions whose veracity is hard to ascertain. For example, discrete fracture networks are routinely represented by abstract fractal models such as a Sierpinski lattice (e.g., Doughty and Karasaki, 2002) and the Watanabe and Takahashi (1995) structures. Numerical simulations of transport processes on such networks (e.g., Roubinet et al., 2010, 2013; Gisladottir et al., 2016) employ analytical solutions that make multiple assumptions about both the network topology and the underlying physics (e.g., Tang et al., 1981; Sudicky and Frind, 1982; Ruiz Martinez et al., 2013).

In materials such as shale, granite and other crystalline rock, fractures form the main pathways for flow due to the high contrast in permeability between the fracture and matrix. The transport of chemical species primarily occurs along these flow pathways in fracture networks that also facilitate fracture-matrix species exchange that is governed by diffusion. Continuum methods that use averaged parameters such as permeability and porosity cannot resolve the flow in the fractures and thus are also unable to capture the transport mechanisms caused by flow in these fractures along with the fracture-matrix interactions. The DFN approach, where fractures are represented as 2D planes in 3D space, is able to overcome these drawbacks of classical continuum approaches. A DFN is created using fracture stochastics representations based on parameters such as size, orientation, apertures and locations of fractures, and information from field characterization at a given site. A mesh of the network is generated, and then flow and transport solvers are applied to the mesh. The same continuum flow and transport solvers can be used for the DFN approach by ensuring the appropriate permeabilities are used. Typically, fracture permeabilities are evaluated using parallel plate approximation otherwise known as the ‘cubic law’ (see section 3.2.1) Although the fidelity of the physics is fairly improved by the DFN approach, they can be computationally expensive because the stochastic nature of the DFN approach, requires the performance of uncertainty calculations on several hundreds to thousands
of realizations a typical DFN simulation takes 50-100 CPU hours. Such uncertainty calculations on a realistic site can be a daunting and time consuming task. This strongly suggests that reduced models are needed that are faster and somewhat accurate in comparison to a DFN calculation.

5.1.2. Flow and transport in fractured rocks: ML and graph-based models for UQ

Workflows for characterization of sub-surface fracture systems lend themselves very well to the inclusion of ML algorithms (Viswanathan et al., 2018; Bergen et al, 2019). The efficacy of an ML algorithm in predictive scenarios depends upon whether the algorithm has previously seen a similar input feature set, and requires a rich training dataset. Additionally, ML algorithms must have a sufficiently broad validation data set to determine whether predictions are of acceptable accuracy. For example, observed solute breakthrough times at a field site could be compared to the ML emulator output. Such large datasets are now openly available because of the confluence of improvements in sensors that enable high-resolution measurements at a higher sampling rate (see Section 2), advances in high-performance computing that facilitate the generation of vast amounts of large scale simulation data, and improvements in infrastructure to store and share data over the cloud (Karpatne et al, 2018). Numerous ML algorithms have been developed that cater to analysis of networks, in application areas like cybersecurity, search engines and social media. These algorithms have been adapted seamlessly for physical science domains, including fracture networks. Of foremost importance is the ability to consider a fracture network in its entirety and reduce the data through features other than size, as often, the key characteristics determining the fate of transport in fracture networks relates to the topology rather than fracture lengths and apertures. The analysis of subsurface fractured systems thus benefits immensely from principles drawn from a combination of graph theory along with ML, to aid identification of a lower dimension manifold that includes the physical, hydrological and topological features (Srinivasan et al, 2018) of the fracture network. A further benefit of an accelerated ML-driven workflow is the ability to incorporate exploration of uncertainty space for characterizing system uncertainties.

One can reduce the complexity of fractured systems by using graph theory approaches to identify key sub-networks that form the backbones for flow using ML (Viswanathan et al., 2018). Figure
11 is a conceptual diagram that shows how to map a DFN to a graph to reduce the degrees of freedom. By solving for flow and transport on the graph, the computational efficiency is greatly increased but the solution is not as accurate. Machine learning can be used to correct the graph solution to match the DFN (Srinivasan et al., 2018).

These graph-based approaches can also be used to identify the dominant flow path through a fracture network. The premise behind these approaches is based on the observation that the entire fracture network does not contribute to the flow and transport, but only a subset or preferential (Abelin et al., 1985; Tsang and Neretnieks 1998) pathways that lead to flow channeling. One first maps a DFN to an equivalent graph, then identifies the dominant flow path, using ML based reduction methods that use topological features such as connectivity (Valera et al., 2018) or together combined with hydrological features such as mass flux that have been shown to improve the quality of the backbones (Srinivasan et al., 2019). Once these subnetworks in graph space are identified, they are then mapped back to the DFN space. Flow and transport calculations are then performed on these subnetworks in the DFN space that reduce the computational time by 50%, and yet give accurate results (Hyman et al., 2017; Hyman et al., 2018).

5.1.3. Potential for coupled T-H-M-C processes within a UQ framework

The previous two subsections describe a combination of simplified stochastic models and graph-based ML emulators that enable robust uncertainty quantification for flow and transport in fractured systems. Unlike high-fidelity modeling (section 4) where great progress has been made in incorporating T-H-M-C effects on flow and transport, limited uncertainty quantification work exists in this area. There is a great need for including UQ and T-H-M-C under the same framework for fractured systems because coupled processes can significantly alter the fracture volume fraction which in turn can cause major changes in flow and transport through a fracture network (because of alterations to the fracture network geometry but not necessarily the topology). Including T-H-M-C greatly increases the parameter and model uncertainty in predictions as well. Progress has been limited because of the difficulties in the formulation of simplified models or emulators for the complex coupled processes. Here we identify specific application areas with the need for T-H-M-C processes and suggest a path forward.
Additional uncertainties are inherent in the simulation of coupled T-H-M-C processes compared to flow and transport in a fracture network alone. As more processes are considered, the number of parameters involved in the simulation increases. Unlike in the case of contaminant transport, where the prediction of uncertainties for field variables such as concentrations at specific locations and times is important, T-H-M-C coupled processes are usually more focused on system-scale performance metrics or in the probability that a particular (typically adverse) self-organized runaway phenomenon may occur. As with flow and transport processes, the challenge then is to improve our understanding of the minimum number of stochastic realizations needed to simulate and to quantify the uncertainty associated with relevant performance metrics. Next, we describe a few examples of T-H-M-C problems in key applications that merit uncertainty quantification:

1) **Shear stimulation of a fracture network in an EGS system and propagation of injection induced seismicity (IIS):** Shear mechanical failure of a connected network of fractures can either lead to enhancement of permeability that improves circulation as in the case of EGS, or to undesirable and excessive pressure diffusion leading to IIS (Ellsworth 2013; Nicholson and Wesson 1992; Weingarten et al., 2015). The main factors that control shear failure of fractures are the stress field, pressure increments that result from injection, the orientation of fractures (typically failure occurs only within a very narrow range of “favorable orientations” (see section 2)), and the friction coefficient of the fractures. Whereas unfavorably oriented fractures are unlikely to experience shear failure, they may still facilitate rapid pressure diffusion to connected fractures that are favorably oriented. Clearly, there are uncertainties in all these factors, including uncertainties in estimates of the initial stress field. *In situ* monitoring of stimulation operations as described in Sections 2.2.2 and 2.3 could provide a Bayesian data assimilation or ML framework to evaluate these uncertainties and to refine estimates of fracture orientation distributions, *in situ* stress and fracture properties. In connection with IIS, it is important to note that various modeling approaches have been developed to improve understanding of IIS – ranging from equivalent continuum models based on the uncoupled groundwater flow (pressure diffusion) equation, fully poroelastic models, and partially coupled models that include representations of discrete fractures. Most field-scale model applications employ ECM models (see section), which
are calibrated to explain IIS occurrences \emph{a posteriori}. Many of these models employ unrealistically large values of hydraulic diffusivities (Brown et al., 2017; Shirzai et al., 2016; Kernanen et al., 2014; Langenbruch et al., 2018), which are calibrated to match IIS occurrence. It is questionable whether these calibrated models that contain inadequate physics, are suited for projecting evolving seismicity at problematic sites.

Several recent modeling studies have shown that the propagation of seismic slip in fractured rock involves an apparent diffusivity (“seismic diffusivity”) that can exceed the hydraulic diffusivity by more than an order of magnitude (Zareidarmiyan et al., 2021; Haagenson and Rajaram 2021). Given the uncertainty associated at various levels with modeling IIS including data assimilation, T-H-M-C process models and the locations of existing faults, a general-purpose probabilistic risk assessment framework is needed (Baisch et al., 2019; Bommer et al., 2015). Such a framework can incorporate reduced-ordered models of the T-H-M-C processes to reduce the computational burden or augment the high-fidelity models with the reduced-ordered models. These reduced-ordered models can be ML-based or graph-theory based similar to the ones discussed in Section 5.2.2 or a combination. ML models can be built by training on data that is generated with high-fidelity T-H-M-C models for ranges of input parameters such as injection flow rates, background stresses, fracture properties, and by recording the output quantities of interest such as slip. Approaches such as deep neural networks can be used to train on these input/output quantities. Such ML models can then be incorporated in risk assessment frameworks that use UQ approaches (Tartakovsky 2016). Along the lines of graph-theory approach, for incorporating T-H-M-C effects, dynamical graph approaches could be used to evaluate fracture evolution (Mudunuru et al., 2019; Hunter et al., 2019) and the resulting mechanical system variables such as displacements and stresses can be used to update flow and transport parameters that are affected by mechanical variables such as permeability and porosity. These parameters can then be updated in a graph-based flow and transport model. This type of full T-H-M-C graph-based approach development is a potential area of research. The risk assessment frameworks with these emulators can be used to address questions such as: can flow rate or injection pressures be modified or sections of the injection well screen be sealed off to mitigate IIS risks during deep well disposal? Uncertainties in location/existence of major faults may be incorporated into such a framework based on structural geological constraints as well.
2) Formation of short-circuits and permeability management in EGS systems: As described in Section 4, short-circuits refer to the formation of undesirable dominant preferential flow pathways in EGS systems, that leads to reduced energy production rates. Short-circuits sometimes occur during or soon after stimulation processes (Parker 1999). In most cases, short-circuits are attributed to thermo-mechanical effects (Kohl et al., 1995; Hicks et al., 1996; Taron and Elsworth 2009; Fu et al., 2016), that cause the matrix blocks to contract because of cooling. This contraction results in enlargement of the fracture apertures which leads to an increase in permeability and to cooler than normal production temperatures. At even longer time scales, precipitation-dissolution cycles influence fracture permeability evolution. In some instances, precipitation can limit short-circuits by counteracting fracture enlargement. Conversely, precipitation can reduce permeability and energy production even in an ideally functioning EGS system that does not experience thermo-mechanical short-circuiting. A-priori estimates of the occurrence and severity of short-circuits and permeability loss are of great value in the design, operation and sustainable management of EGS. Management frameworks that can adaptively identify optimal operational strategies to remedy short-circuits and permeability loss are of great interest. The major uncertainties in the evolution of EGS fracture systems put both of these endeavors in the class of “decision making under uncertainty”. A systematic UQ framework is an essential component of an adaptive management system for EGS. The graph-theory and ML approaches described in the previous section are uniquely suited to complex fracture networks. While these works focused on the first breakthrough, there are exciting and promising possibilities for orienting these approaches towards the performance metrics of interest in EGS.

3) Thermally-Controlled Mineral Reactions in High-Level Nuclear Waste Repositories: Another area of interest is to determine the risk of radionuclide release outside of a specified compliance boundary. The risk metrics of interests are similar to a contaminant transport problem and include first arrival time and maximum concentration. The simplified models and emulators approaches are well suited for simulating these quantities of interest. However, because radioactive waste disposal sites must be assessed over 1000s of years, the underlying uncertainties are controlled by complex thermo-hydrologic flow regimes in fractured media, phase change and boiling regimes, and by chemical alteration of mineral phases within the rock.
matrix, which influence radionuclide immobilization (Viswanathan et al., 1998). While the underlying fracture network is a major source of uncertainty, the thermo-hydrologic controls are likely quite robust. In the case of energy transport in fractured rock, the rock matrix participates actively and the action is not restricted to fractures. As a result, the large-scale thermal regime resulting from repository heat generation is likely to be insensitive to the underlying fracture network structure, although the dominant flow paths will no doubt be influenced by the fracture network. A UQ framework for this problem needs to quantify the uncertainty in the advective velocity field and its evolution. Subsequently, radionuclide transport needs to be considered because T-H-M-C processes dynamically change the fracture network dynamically. Since simplified models and emulators cannot currently handle these complex T-H-M-C processes, one-dimensional T-H-M-C simulations along streamlines may be a good first step to developing computationally efficient T-H-M-C simulations that can be incorporated into a Monte Carlo UQ framework (Viswanathan et al., 2004). In addition, graph-based flow and transport approaches (Karra et al., 2018) can be extended to include reactions and to serve as fast emulators for thermo-hydro-chemical process models. ML models inferred from high-fidelity simulations is an alternate emulator approach. Emulator approaches with UQ settings, such as Monte Carlo, can be used to evaluate the performance of waste disposal strategies.

4) Stress and Mineral Reactions in CO₂ sequestration and EGS-CO₂ systems: For EGS and CO₂ sequestration systems, some of the performance metrics, such as injectivity, potential for caprock and wellbore failure, depend on chemical and mechanical feedbacks associated with the fracture network. Mineral precipitation and dissolution reactions can open and close fracture apertures, thus affecting these metrics. Injecting large amounts of CO₂ over large periods of time can also lead to large mechanical stresses that also affect these metrics. Therefore, performance assessment of CO₂ sequestration sites must include T-H-M-C processes to improve quantification of injectivity and leakage. Current state of the art assessment of a site combines multiple realizations of simplified models with a few detailed T-H-M-C simulations to qualitatively assess the site (Stauffer et al., 2009; Viswanathan et al., 2008). Approaches taken by the National Risk Assessment Program for Carbon Sequestration use an integrated assessment model that links surrogate models together for performance assessment. These surrogate models are computationally efficient emulators of physics-based models for processes such as
inj ectivity, wellbore leakage, caprock leakage and induced seismicity (Vasylkivska et al., 2021; Jordon et al., 2015; Chen et al., 2018; White and Foxall 2016). This framework enables T-H-M-C processes to be incorporated into the integrated assessment framework if the appropriate surrogate model can be developed. Again graph-based or machine learning emulators are a promising approach to incorporate T-H-M-C feedbacks into a surrogate model for wellbore and caprock leakage.

Physics-informed ML-based reduced-ordered models are potentially a good approach to account for T-H-M-C effects for the above discussed subsurface applications (Ahmmed et al., 2021; Chen et al., 2020; Mudunuru et al., 2017). These ML approaches can use either supervised or unsupervised methods but a good understanding of the input variables and the quantities of interest is required. The data to train these models can be generated by running Monte Carlo based realizations of the system within the bounds of the known ranges for the T-H-M-C parameters. The derived ML models can be used as surrogates in UQ frameworks to either evaluate the uncertainty to identify where new observations are needed or to evaluate risk (e.g., in the case of IIS scenarios). However, care must be taken when using ML models because these ML models do not perform well beyond the limits of the training data. If the input or QOIs (quantities of interest) change, a ML model would have to be re-trained or a new ML would need to be built.

5.1.4. Uncertainty quantification methods

The discussion above makes it clear that robust UQ for predictions of flow and transport in fractured rocks must address both model and parametric uncertainties, with the former comprising two parts: uncertainty in the fracture network representation and uncertainty in the representation of the physics, chemistry, geology, etc. A vast body of literature has examined various facets of predictive uncertainty, especially in the context of nuclear waste repositories. Common sense dictates that no mathematical model can predict point-wise quantities, such as a solute concentration at a particular location in the subsurface, with any reasonable degree of certainty. On the other hand, science-based predictions of integral characteristics, such as the amount of water stored in a fractured formation or the first arrival time of a solute, are expected to be accompanied by confidence intervals derived from UQ.
Figure 11: Workflow of discrete fracture network model reduction approaches.

It might be impossible and, from a practical perspective, unnecessary to untangle various sources of predictive uncertainty: alternative models are alternatively parameterized, and predictions are made in the ensemble sense, i.e., averaged over multiple realizations. With this caveat, results from Monte Carlo simulations suggest that model uncertainty--expressed as the choice between the stochastic continuum representation and discrete fracture network models--is relatively unimportant in predictions of the integral quantities of interest (Finsterle, 2000; Selroos et al., 2002). Monte Carlo simulations provide a means for uncertainty quantification and also for the probabilistic identification of the backbone of a discrete fracture network (Osthus et al., 2020) and the construction of simplified representative models of such networks (Geier et al., 2019).

Standard Monte Carlo simulations are conceptually straightforward, robust, and trivially parallelizable. Yet, a large number of realizations (forward model runs) are required for the simulations to converge, which makes them impractical for large models. The need to improve this slow convergence rate drives the field of UQ. Multilevel/multifidelity Monte Carlo can speedup standard Monte Carlo by orders of magnitude (e.g., Berrone et al., 2018; 2020; O’Malley et al., 2018), although such a performance is not guaranteed if the goal is to compute...
the full distribution of a quantity of interest rather than its mean and variance (Taverniers and Tartakovsyz, 2020; Taverniers et al., 2020).

Various types of polynomial chaos expansion, PCE, (e.g., Agada et al., 2017; Koohbor et al., 2019) are often used as a computationally efficient alternative to Monte Carlo simulations. They serve as robust surrogates for ensemble-based computations, including inverse modeling, data assimilation, and global sensitivity analysis (Wang et al., 2017; Zoccarato et al., 2020). The efficiency of a PCE decreases with the number of random parameters (aka “stochastic dimension”) used to characterize a (fractured) subsurface environment: the so-called “curse of dimensionality” refers to the threshold beyond which PCE becomes less efficient than Monte Carlo. Polynomial chaos uses truncated Karhunen-Loeve expansions (or their analogues) to approximate a spatially varying random parameter representative of heterogeneous environments. The number of terms in a PCE, i.e., the stochastic dimension, increases as the correlation length of the parameter decreases. This renders PCEs ill-suited for poorly correlated heterogeneous fractured rocks.

State-of-the-art UQ techniques have their strengths and limitations. The complexity of most contemporary models, e.g., their multiscale and multiphysics nature, argues for the deployment of non-intrusive UQ strategies, such as various types of Monte Carlo simulations and stochastic collocation on sparse grids (Berrone et al., 2017). Unlike their intrusive counterparts (e.g., the PCE-based stochastic finite elements), such sampling techniques allow for the use of existing simulation software and, hence, are more likely to be embraced by the community. On the other hand, the intrusive methods are more robust, which illustrates yet another facet of the complexity versus functionality trade-off.

5.2. What we learned and path forward

Most of the state-of-the-art UQ techniques have been deployed to analyze various sources of uncertainty in models of flow and transport in fractured rock. As discussed above, the goal of such techniques is to outperform standard Monte Carlo simulations in terms of computational efficiency. For non-intrusive (i.e., sampling-based) methods, this aim is tantamount to finding a
method that requires the smallest number of realizations to converge. In other words, the need to replace Monte Carlo arises only when a model’s forward solve is computationally expensive. An accurate and robust surrogate/emulator, which carries virtually no cost to execute, obviates this problem. However, the training of such emulators requires a large number of realizations as well.

In the past decade, advances in machine learning techniques, the ease of access to heterogeneous computing with GPUs and TPUs via cloud, and the availability of machine learning toolkits such TensorFlow, PyTorch, etc., has led to a significant increase in the use of machine learning in workflows of domain experts including subsurface scientists. To ensure that the governing physics for the subsurface flow and transport processes are satisfied, several physics-informed and physics-constrained approaches have been developed (Tartakovsky et al., 2020; He et al., 2020; Mudunuru and Karra, 2021). These approaches can be used directly to assimilate data to characterize parameters such as permeability (Tang and Durlofsky, 2020; Wu and Qiao, 2021) and to quantify uncertainty (Yang et al 2021). Alternatively, fast-running emulators (4-5 orders of magnitude faster than high-fidelity simulations) can be built with these approaches that mimic the physics reasonably well, within the bounds of the trained data. These emulators can then be employed in UQ, parameter estimation and data assimilation workflows. Although most of these approaches have been applied with two- and 3D continuum porous media problems, they can be extended to workflows involving fractured systems. For instance, Mudunuru et al., 2020, recently proposed an approach that combines both high-fidelity simulations and emulators for hydrocarbon production in a shale system.

Graph-based emulators have been recently proposed as an alternative to performing physics calculations on complex meshes such as DFNs. Karra et al., 2019 performed direct flow and transport calculations on graph equivalents to DFNs. In their approach, a fracture is reduced to a node on the graph. Equivalent parameters such as permeability are assigned to the edges of the graph. The equivalent permeability is a function of the geometry of the two intersecting fractures and the distance between the centers of the fractures. The graph can then be treated as a ‘coarse’ mesh, and flow and transport are then solved on this graph. Karra et al., 2019 solved this for steady single-phase flow and compared the transport breakthrough against DFN predictions. Because of the reduction in fidelity, the pressure gradients were found to be consistently under-
predicted, which led to slower breakthrough times. To reduce this error, they performed a correction that used machine learning that resulted in reasonably accurate predictions relative to other models. A greater than four orders of magnitude computational speedup was observed with the graph approach, making it amenable for uncertainty quantification calculations.

These graph-based emulators have been extended to transient single-phase flow (Srinivasan et al., 2020), and were used to estimate production rates for oil & gas application (Dana et al., 2020). Although these graph-based emulators focus on flow and chemical transport, a similar methodology can be used to emulate coupled flow and heat. However, this will require incorporation of components that capture heat exchange between fractures and matrix as well as a source/sink term lumped term for incorporating the exchange. In the realm of fracture mechanics, there have been several works that use graph emulators as well. These emulators have been used to perform fracture propagation calculations with significantly smaller computational cost (Khodabakhshi et al., 2019; Mudunuru et al., 2019; Hunter et al., 2019). Although all these works related to graph emulators in the area of fractured systems show that these emulators have immense potential to replace high-fidelity models for UQ workflows, there are yet some challenges that need to be addressed. The main one being that the graph emulators need a correction to account for loss in fidelity. ML can play a critical role in bridging this gap. However, this will require the generation of data via large amounts of high-fidelity simulations to train the correction models, which may not always be feasible. Furthermore, the correction may not be generalizable, and the ML method will need to be re-trained for any new conditions. This challenge compounds when coupled T-H-M-C processes are considered, as the interactions are fairly complex and non-linear in nature. Nevertheless, graph emulators with ML-based error correction methods offer a good balance between accuracy and computational cost.

Although the recent advances in computational resources and tools have led to an increase in the fidelity of the models, that is, adding more resolution and more physics to the models, there is still a need to incorporate and to account for uncertainties that exist because of the lack of measurements (leading to uncertainty in fracture properties and model parameters) and the lack of understanding of complex coupled process models. Even if there is a reasonable amount of field data, e.g., fracture properties, the values are only known in a statistical sense. For all these
reasons, methods that quantify uncertainties in a statistical sense are needed but they typically require running the underlying physics-based models several hundreds of times. Furthermore, subsurface applications involving injection of fluids such as EGS and CO₂ sequestration pose the threat of creating seismic events, and also the possibility of deadly leaks through existing fissures and faults. These applications thus require an additional layer of calculations that account for various scenarios to quantify risk. Despite the availability of HPC and massively parallel high-fidelity codes, running hundreds of realizations is still computationally intractable and has motivated the evolution of significantly faster emulators and reduced-ordered models that mimic these high-fidelity models to variable extents. Recently, emulators based on graphs have shown promise with a speed of over 4 orders of magnitude with reasonably accurate results. Emulators have been used for coupled flow and transport, and fracture mechanics applications, separately. There is the potential to extend these methods to coupled T-H-M-C processes for various subsurface applications. Advances in machine learning and the availability of several open ML tools has also led to the incorporation of these methods in physics-based workflows including subsurface fracture media applications. More recently, unsupervised methods such as matrix methods and supervised methods such as CNNs have been used to build emulators and reduced-ordered models for several subsurface applications. Immediate areas for research and development investments is in the use of deep neural networks to build towards T-H-M-C emulators and to use unsupervised methods for understanding the complex coupling of various physical and chemical processes. Physics informed and constrained ML methods can be used for parameter estimation and UQ from field and laboratory measurements. In addition, these methods, when coupled with methods such as active learning, can provide insights into where measurements need to be performed or sensors need to be placed. Thus, ML can serve as a powerful tool to bridge the gap between modeling and measurements either in the laboratory or in the field.

6. Conclusions

Demystifying the science of flow and transport in subsurface fractured systems is crucial to solving several grand challenges in the fields of energy, water and global security. While our understanding of fractured rocks has advanced significantly in the last forty years (Section 6.1), particularly in terms of the effects of coupled thermo-hydro-mechanical processes (T-H-M-C) on
flow and transport, several important challenges remain. It is now abundantly clear that integration of laboratory and field experimentation, physics-based models and their surrogates, and sensitivity and uncertainty quantification techniques is needed to move the field forward (Section 6.2).

6.1. Recent advances

A significant component of fractured rock research must occur at the field scale. Field studies provide understanding and a means to test technology under *in situ* conditions, across scales, in natural geologic environments with their heterogeneities, and in the presence of natural driving forces. Experience of the past two decades has shown that highly characterized and densely monitored sites dedicated to long-term observations are critical for advancement of fractured rock research. The latter has benefited from transformational advances in monitoring technologies, such as the development of integrated borehole tools for coupled processes, the advent of real-time geophysics, and the emergence of fiber optics technologies for high-resolution spatial monitoring. While the characterization and monitoring tools deployed at these sites might not be feasible or cost-efficient in most subsurface applications, these dedicated observatories provide high-resolution data sets that (1) serve to elucidate the critical T-H-M-C mechanisms controlling fracture flow and transport and (2) aid in the development and validation of predictive models. Forty years ago, only a few dedicated *in situ* sites focused on fractured rock behavior. The past two decades have witnessed many new facilities devoted to interrogation of fractured systems across multiple scales and characterization of a variety of geological environments. These facilities are often utilized in open team settings, with researchers from multiple organizations and various disciplines working together to contribute to model validation. These subsurface facilities provide opportunities to conduct cutting-edge research on fractured rock, to field-test new monitoring techniques, and to explore novel ways of utilizing the subsurface, spanning fundamental fracture hydrology and mechanics research applications such as geothermal energy, fossil energy, CO₂ and energy storage, nuclear non-proliferation and radioactive waste disposal. Emerging areas such as hydrogen storage and blue hydrogen, which are needed to transition away from fossil fuels, will also utilize the fractured subsurface.
Laboratory experiments are used to determine the fundamental T-H-M-C controls on fracture flow and transport. Experimental data and observations inform interpretations of field observations and, more importantly, provide a mechanism for validating numerical simulations of fracture processes under controlled initial and boundary conditions. Experimental studies show that fracture roughness, aperture, and contact area control fracture permeability and point the way to the still elusive geometry-based model for the prediction of flow in fractures. Today’s laboratory experiments incorporate x-ray and optical methods to image dynamic interactions among fractures, stress, and flow processes under replicated subsurface conditions to provide insights into flow and transport under realistic conditions encountered during subsurface activities. Experiments on coupled T-H-M-C processes have quantified the impact of normal stress on permeability, shown that fracture reactivation can increase or decrease permeability depending on displacement and rock mineralogy, and demonstrated the complex, variable impacts of dissolution on permeability, all of which are critical feedbacks in field applications. With the advent of machine learning technology and ever-increasing complexity in simulations, laboratory studies provide fundamental constraints to the properties and behavior of massive fracture networks.

The goal of physics-based numerical simulations is to understand, predict and optimize the behavior of a site-specific fractured system. These simulations are developed from model components that are validated by high-fidelity simulation of fracture experiments (Section 3), proved on field experiments (such as the EGS Collab project and the Mont Terri underground research laboratory described in Section 2) and informed by site characterization of stresses, geology, rock properties, fracture distribution, and fracture characteristics. The required fidelity of numerical simulations of fractured rock depends on the complexity of the fractured system being modeled and the amount of characterization data available. In many cases, it will require the latest in high-fidelity simulations that rely on high-performance computing resources and state-of-the-art meshing algorithms to represent the underlying fracture network and the relevant coupled processes. For example, the international DECOVALEX initiative has demonstrated how the integration of dedicated field experiments with coupled processes and high-fidelity simulations leads to new scientific advances, guides the development of new simulation techniques, and helps reduce conceptual model uncertainty via comparative evaluation.
However, for many fractured subsurface systems, little information about the underlying fracture network is available, rendering uncertainty quantification critical. Since the fracture network structure is often a critical feature that controls flow and transport through the subsurface, discrete models that explicitly represent fractures might yield higher fidelity in mechanistic predictions of the physics and chemistry of the system. Since fracture locations are not known with certainty, these discrete fracture models need to be combined with uncertainty quantification techniques that can explore the large parameter space of a typical field-scale fractured site. Continuum models may be appealing because they require fewer parameters (less extensive field characterization) and are computationally more efficient than discrete models. However, continuum models make underlying assumptions about the fracture network and do not allow the modeler to test hypotheses on the structure. For many coupled T-H-M-C fracture flow problems, this lack of structure can lead to inaccurate results (Section 4). Recent hybrid methods that employ graph-based machine learning emulators for discrete fracture models are a promising approach to accelerate discrete models. This allows them to be placed within a robust uncertainty quantification framework (Section 5). Emulators linked to physics-based models enable fast exploration of a large parameter space. These methods have primarily been used for single-phase flow without incorporation of T-H-M-C processes, but there is a path forward to expanding these methods for more complex physics using methods such as streamline modeling combined with graph theory.

6.2. Path forward

With the latest advances in laboratory and field experiments, physics-based models and uncertainty quantification techniques, we are well positioned to make significant progress on understanding, predicting and controlling subsurface fractured systems in the next decade. To move the field forward we discuss how to tackle the next frontiers, attempting to answer the following questions: (a) What new monitoring technology is needed and what are the emerging scientific questions to be tackled with field experiments? (b) What laboratory experiments are needed to fill critical knowledge gaps? (c) What is the next step in advancing high-resolution coupled process models? (d) What constitutes a pragmatic and efficient hybrid modeling architecture and what surrogate modeling approaches advance fundamental understanding of T-H-M-C processes in fractured rocks?
Field Observations: Technologies such as 3D downhole probes and distributed multimodal optical fibers are evolving quickly and offer the possibility to track signals over spatial scales from localized µm-scale to distributed measurements over km-deep boreholes while potentially spanning large time scales. These techniques have the potential to finally probe complex phenomena, such as long-term creep and sealing processes, that can greatly affect fracture flow over large time-scales. More progress is also needed to constrain geomechanical parameters such as stress, fractured rock elasticity, and plasticity. The ability to monitor chemical changes in distant fractures is elusive, but the development of distributed chemical fiber-optic based sensing (DCS) is a promising path to complement existing DTS-DSS-DAS fiber deployments and improve characterization of chemical processes. Instruments for long term deployments must be hardened to withstand the harsh temperature and chemical environments often experienced in deep subsurface systems. Natural analogs that probe geologic conditions operating over geologic times are another promising avenue for characterizing longer-term chemical processes.

Given the difficulty and expense of drilling, the characterization and exploration of deep subsurface processes has historically been data-scarce. With new developments in high-resolution time-lapse sensing, geoscientists are dealing with increasingly large data sets. This will require better data collection and archiving methods combined with advances in techniques for joint assessment and inversion of multiple parameters at the same location, such as combining seismic waves, strain measurements, pore pressure, and fluid chemistry data. Already, technologies such as DAS can generate upwards of 10s of TB/day. Advanced machine learning techniques are already available for data reduction and analysis. Further improvements should enable “ultra-fast” big data analysis and interpretation, enabling both new scientific breakthroughs and adaptive near real-time control of subsurface operations.

Laboratory Measurements: Experiments that characterize the full fracture geometry should be developed and used to validate fundamental flow simulations. This same effort provides a path forward to the intercomparison of permeability measurements and the development of a predictive framework for the permeability of subsurface fractures. Experiments should include studies of simple and complex networks that better inform field observations and large-scale modeling where network flow and transport play a dominant role. A major difference between single fractures and fracture networks is the presence of fracture intersections that enable the
formation of 3D dominant flow paths that break the universal scaling observed in 2D single-fracture behavior. A great deal of laboratory research has been conducted on hydro-mechanical processes for single fractures but more research is needed on the hydro-mechanical coupling in the presence of fracture sets and fracture networks. In addition, more research is needed on the effects of fracture reactivation and chemistry on flow, transport and deformation. Chemo-mechanical coupling is underappreciated but important in many subsurface applications that requires a better understanding of reaction halos on the long term permeability of fracture networks under stress. More laboratory studies of fracture network formation are needed to provide insight into processes that control the formation of specific network topologies and how intersections are formed to provide constraints on statistical representations of fracture networks for specific subsurface geologic and tectonic conditions. Permeability studies should also consider solute transport characterization, both for model validation and for development of methods to characterize subsurface fracture area. These are challenging experiments that are needed to assess problems such as short-term prediction of radionuclide migration for nuclear nonproliferation and longer-term predictions for problems such as carbon sequestration, geothermal energy and radioactive waste disposal.

*Physics-based Modeling:* Simulators that take advantage of the latest computer architectures and algorithmic developments have made great progress in representing fractures as discrete features embedded in a rock matrix continuum. DFM models are well positioned to simulate T-H-M-C processes with the real possibility of modeling large complex 3D fractured rock masses. The key challenge is to constrain these simulations using multiple data streams from field observations and laboratory experiments. An additional computational challenge arises from the extensive multiscale heterogeneity within fractures and faults that can often control system behavior. Flow at these multiple scales results in preferential pathways that often dominate fractured system behavior. In the past, resolving flow within a single length scale was so computationally demanding that it required using high performance computing resources. Today, modern DFN and DFM allow for multiple length scales of heterogeneity to be represented within field-scale models using a desktop machine. Key features such as aperture variability and surface roughness can now be included into large network simulations to determine the relative importance of single fracture and network processes. Although rigorously including fracture growth due to
geomechanics is still out of reach, relatively new kinematic DFNs mimic the growth of fractures using simplified rules of nucleation, growth, and arrest. While lower-fidelity than a full geomechanics-based simulation, which remains infeasible for thousands of fractures, these models create fracture networks with more physically realistic structures than a purely stochastic network.

Uncertainty Quantification: Despite improved characterization and physics-based models, subsurface fractured systems are and will always be plagued with uncertainty in both parameters to characterize a particular system and our ability to observe relevant quantities of interest. Future modeling efforts are likely to incorporate a pragmatic and efficient hybrid modeling architecture that combines physics-based models and their surrogates. The physics-based models must be validated against field observations and laboratory experiments to ensure key trends in observations are simulated and the quantities of interest are properly history-matched and predicted. The surrogate models can accelerate the physics-based computations to allow for full interrogation of the uncertain parameter space of a typical fractured site. The choice of which surrogate to use and the level of fidelity of the physics-based model should be guided by data availability, the quantities of interest, and the logistics of the decision-making process.

This brief but not comprehensive review illuminates the progress made in the last 40 years related to the behavior of fractured porous media. The fracture community is on the cusp of great advances in understanding coupled processes in fractured rock in the next decades because of the simultaneous and continuous advancements in field testing, laboratory experiments, physics-based modeling and uncertainty quantification in the past decades. We hope this manuscript will focus talent and effort on needed future directions to achieve this advancement to aid improvements in our ability to work and interact with the Earth’s subsurface in safe and sustainable ways.

7. Acknowledgements

HV, JWC, JDH, SK and GS gratefully acknowledges support from the Department of Energy, Office of Science, Office of Basic Energy Sciences, Geoscience Research program under Award
Number (LANLE3W1). DT thanks the US Department of Energy GTO Award Number DE-EE3.1.8.1 “Cloud Fusion of Big Data and Multi-Physics Models using Machine Learning for Discovery, Exploration and Development of Hidden Geothermal Resources” for support of this work. HR thanks the National Science Foundation Hazards SEES project EAR 1520846 at the University of Colorado Boulder. LJPN acknowledges support of this work by the U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences, Geosciences Research Program under Award Number (DE-FG02-09ER16022). Funding for JB and YG on this review paper was provided by the U.S. Department of Energy’s Office of Fossil Energy and the U.S. Department of Energy’s Geothermal Technologies Office under contract DEAC02-05CH11231. J.A-F. was funded by the Office of Energy Efficiency and Renewable Energy, Geothermal Technologies Office, the U.S. Department of Energy (DOE) under Award Number DE-AC02-05CH11231 with Lawrence Berkeley National Laboratory (LBNL). Since this is a review article, no new data is presented in the manuscript.

JB, JAF, and YG wrote the field observation section 2. JWC and LJPN wrote the experimental section 3. JDH, HR and HV wrote the computational modeling section 4. HR, HV, SK, GS and DT wrote the uncertainty quantification section 5. HV integrated all sections and all authors have contributed to ensuring the overall narrative is a coherent story. All authors have no conflict of interest. The authors thank Shaoping Chu for help with formatting references.
8. References


Pyak-Nolte, L. J. (2019). Fracture Specific Stiffness: The critical link between the scaling behavior of hydro-mechanical coupling in fractures and seismic monitoring. In *Science of Carbon Storage in Deep Saline Formations* (pp. 311-335). Elsevier. doi: [https://doi.org/10.1016/B978-0-12-812752-0.00014-9](https://doi.org/10.1016/B978-0-12-812752-0.00014-9)


van Renssen, S. (2020). The hydrogen solution?. *Nature Climate Change, 10*(9), 799-801. doi: [https://doi.org/10.1038/s41558-020-0891-0](https://doi.org/10.1038/s41558-020-0891-0)


