Karst and cave formation in Earth’s upper crust by cooling of CO2-rich geothermal flow

Roi Roded1, Einat Aharonov2, Amos Frumkin2, Nurit Weber3, Boaz Lazar2 and Piotr Szymczak4

1Hydrology and Water Resources, The Hebrew University, Jerusalem, Israel

2Institute of Earth Sciences, The Hebrew University, Jerusalem, Israel

3Department of Earth and Planetary Sciences, The Weizmann Institute of Science, Rehovot, Israel

4Institute of Theoretical Physics, Faculty of Physics, University of Warsaw, Warsaw, Poland

**Abstract**

Porous, fractured rocks in Earth’s upper crust continuously undergo reactive flow, which alters their porosity and permeability. In carbonate aquifers, this process may produce extensive karst, with sizeable cave systems (on the scale of kilometers). Of special interest are “hypogenic” karst and caves formed by upwelling deep-seated flow with no genetic connection to the surface. Despite its vast economic and environmental importance, hypogenic karst formation remains poorly understood.

The present work combines geochemical and numerical analyses with field observations to demonstrate that cooling of CO2-rich geothermal flows and upwelling into a confined permeable layer may induce aggressive karstification and cave-forming processes (“speleogenesis”). As water cools, carbonate solubility increases, inducing undersaturation and forming caves on relatively short timescales (tens of kiloyears). This process explains the location of caves observed in a field case study and captures the characteristics of cave morphology, particularly the formation of the maze-like caves, which are typical of hypogenic karst. From a broad perspective, the results demonstrate that, in conjunction with deep-seated CO2 fluxes, Earth’s geothermal heat loss by upwelling of thermal fluids may extensively shape the upper crust, forming sizeable cave systems.

1. **Introduction**

Large portions of Earth’s upper crust are dominated by carbonate aquifers in which fluid flow transports solute and heat, driving the system out of geochemical equilibrium and inducing fluid-rock interactions1,2. The characteristic high reaction rates of carbonates result in significant void-space alterations, ranging from wide diagenetic variations to the development of karstic aquifers with the formation of extensive cave systems (“speleogenesis”)3,4. Positive and negative feedbacks arise between flow and reaction, fundamentally altering the flow processes themselves5–9. Such feedback is important for a range of different phenomena and applications, such as sustainable management of water resources10,11, geothermal energy use12, CO2 geological storage13,14, and mitigation of induced-seismicity hazards15.

Historically, the formation of karst and speleological systems in carbonates was attributed entirely to the percolation of CO2-enriched meteoric water (“epigenic karst”). In recent decades, however, evidence has shown that karst induced by upwelling of deep-seated flows (“hypogene karst”) is common, and a considerable fraction of known caves are attributed to hypogenic origin. Since these caves form by recharge from below, they are commonly isolated, with no direct genetic link to the surface11,16,17.

Despite the abundance of hypogene processes and their wide hydrogeological implications, the formation process of hypogenic karst remains enigmatic16,18. A variety of dissolution reactions may be involved, mostly encompassing low-pH fluids such as carbonic, sulfuric, or organic acids. Furthermore, various scenarios have been suggested to lead to fluid reactivity, including undersaturation induced by mixing of saturated solutions of different compositions (“mixing corrosion”) and arising because of cross-formational flow, or processes involving condensation of water on rock walls above the water table (“condensation corrosion”). The latter can be highly aggressive if induced by sulfuric vapors19–24.

Additionally, because hypogenic caves form in regions of upwelling of thermal fluids, the cooling of thermal groundwater was suggested to drive dissolution (see, e.g., ref. 25). Dissolution by cooling of thermal groundwater is connected to the retrograde solubility of carbonate minerals, whereby their solubility increases as the water cools. However, this mechanism is thought to play a minor role in speleogenesis and to enhance porosity only in a diffuse and dispersed manner26–28. The present work focuses on this mechanism and shows that previous calculations have underestimated the speleogenetic potential of cooling geothermal fluids in carbonates.

An open question to be resolved is whether the hypogene speleogenesis process is linked to the variety of observed natural cave morphologies and intricate patterns, whose formation details remain a mystery17,28–31. Unraveling the underlying physics and the resulting cave morphologies is a scientific challenge whose resolution should significantly advance our understanding of hypogenic karstification.

From diverse perspectives, this study examines carbonate dissolution driven by cooling geothermal flow: First, a geochemical analysis shows that retrograde solubility leads to highly aggressive karstification. Next, a case study of a group of hypogene caves32 in a confined carbonate aquifer serves as the basis for the development of a conceptual and numerical model, the results of which reveal the optimal conditions for the rapid development of localized dissolution and speleogenesis under cooling geothermal flow in carbonates. Finally, we use a network model to describe dissolution along a bedding horizon and show that the resulting cave morphologies are of a maze-like character typical of hypogenic caves. In a wider context, the findings show that the upwelling of geothermal fluid, in conjunction with deep CO2 fluxes, may extensively shape and karstify carbonate aquifers in the upper crust and form sizable speleological structures.

1. **Conceptual model**

Our conceptual model involves groundwater flow that circulates in deep aquifers (>1 km), where it is heated and considerably enriched with CO222,27,33. Under favorable hydrological conditions, this hot groundwater may upwell through permeable sub-vertical faults and fractures, driven by artesian or tectonic pressures and buoyancy forces18,27. Such upwelling often occurs in a pipe-like manner, either at fault or fracture plane intersections34–37 or at flow conduits that naturally occur on the rough surfaces of fractures38,39. Rapidly ascending fluids will maintain their heat and temperature until they reach a flow barrier (e.g., a shallower aquicludic layer stratum), at which point, beneath the aquiclude, the flow is diverted sideways radially into a permeable bedding horizon of aperture *h* in a soluble layer3,40,41. The radial flow in the horizon cools rapidly by transferring heat to the bedrock below and caprock above. Given the retrograde solubility of carbonates, the rapid cooling significantly increases the calcite solubility product, so the solution becomes undersaturated. In turn, this induces localized dissolution, which over time produces a large cave surrounding the inlet (cave feeder; Fig. 1).

The conceptual model described above is consistent with the configuration and morphological features of the group of caves in our case study. This group comprises dozens of extended caves located along the monocline system of the Syrian arc adjacent to the Dead Sea transform in Israel. Large rooted faults at the base of the monoclines are assumed to facilitate upwelling of thermal flow, as further supported by hydrological and geochemical evidence32,42. The caves are commonly developed along prominent bedding horizons in limestone formations and overlay a thick (>400 m) carbonate succession consisting mainly of massive limestone and dolostone. The cave-forming rock layer is overlaid by low-permeability layers comprising mainly soft chalk and limestone-rich in marl horizons. Speleogenesis patterns typically involve maze-like forms and chamber caves43. The maze caves developed due to a relatively uniform dissolution of the network of conduits formed at the intersection of the horizon and subvertical fracture network.

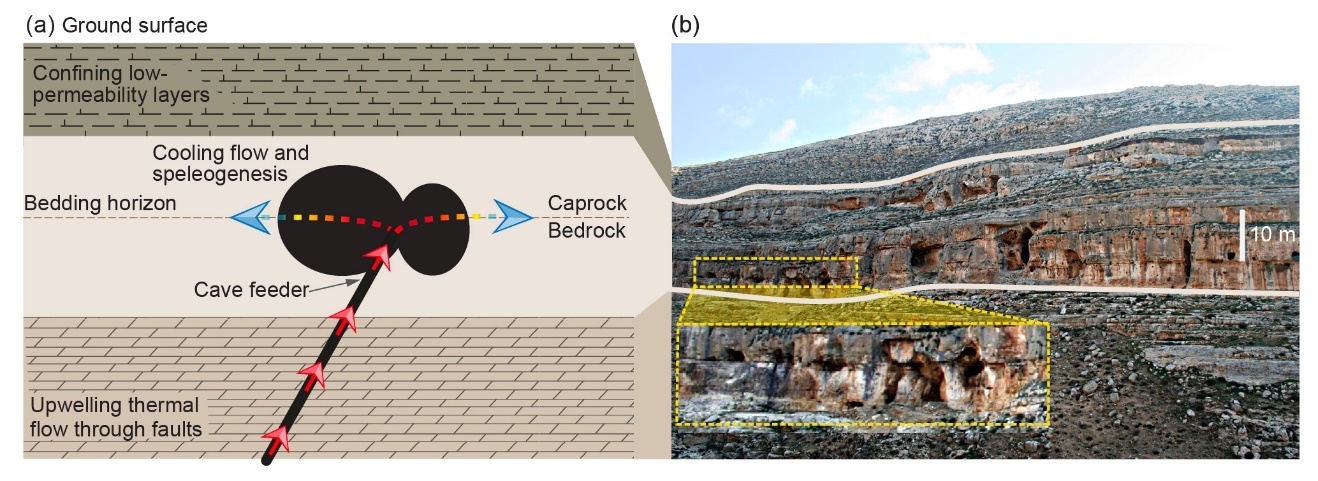


Fig. 1. **a** The conceptual model of thermal speleogenesis and **b** an example from the case study. **a** Thermal, CO2-enriched groundwater upwells through faults and fractures (red arrows) and, upon approaching a confining low-permeability layer, flow is diverted sideways to the permeable bedding horizons in a soluble layer (dashed lines). While the flow through the fault is channeled through the conduit pathway, it retains most of its heat. However, when it is diverted to flow radially along the horizon, it cools rapidly by transferring heat to the rock (gradient color arrows), which leads to a large increase in calcite solubility and solution undersaturation, inducing localized dissolution and speleogenesis (black ellipses). **b** A series of hypogene caves located near Jerusalem, Israel. Beige lines highlight the borders of the main karstified limestone layer, with caves along bedding horizons and fractures (see magnification); the confining low-permeability layer is located above.

The deep-seated groundwater upwelling is assumed to originate from a vast sandstone aquifer because, at present, it also occurs locally (1 to 2 km), although a shallower origin from a carbonate aquifer cannot be excluded (<1 km). Tectonical and hydrological events constrain the speleogenesis duration to a rather large time span on the order of millions of years (Oligocene–early Miocene). Its termination is associated with Neogene uplifting and deepening of the Dead Sea transform, which led to the disconnection of the far-field groundwater flow, water-level drop, and cave dewatering. Further details and a complete description of the caves and the geological setup are available in ref. 32 and references therein.

1. **Geochemical and mathematical analysis**

We use the geochemical PHREEQC code to estimate the dissolution capacity of cooling thermal water (section 4.1). The results of the geochemical analysis are then incorporated into the numerical modeling of (i) heat transport and dissolution in a confined bedding horizon (“Axisymmetric horizon dissolution model,” AHD model), and (ii) investigation of channelized dissolution and cave-pattern formation using a network model. The AHD model is described in section 3.1, and the network model is described in section S5 of the Supporting Information (SI).

* 1. **Axisymmetric horizon dissolution model**

The conceptual model (section 2) is portrayed by a simple physical system (Fig. 2). We consider fluid discharge through a central inlet of radius *r*0, which accounts for the channelized upwelling. The fluid enters radially into a permeable bedding horizon at a constant temperature *T*in and flow rate *Q*. In the horizon, we assume uniform flow, dissolution, and axisymmetry (channelized flow and dissolution are considered in section 4.4.2). Furthermore, assuming orders-of-magnitude contrasts in permeability11,44, flow in the upper caprock and lower bedrock are neglected.

Conductive geothermal heat flux and the initial geothermal gradient are assumed to be negligible compared with the fluid heat input. We further assume that the vertical extent of the domain is relatively large so that the heat transfer to the surface is limited. Otherwise, the cooling rate increases, and dissolution and speleogenesis occur even closer to the inlet.

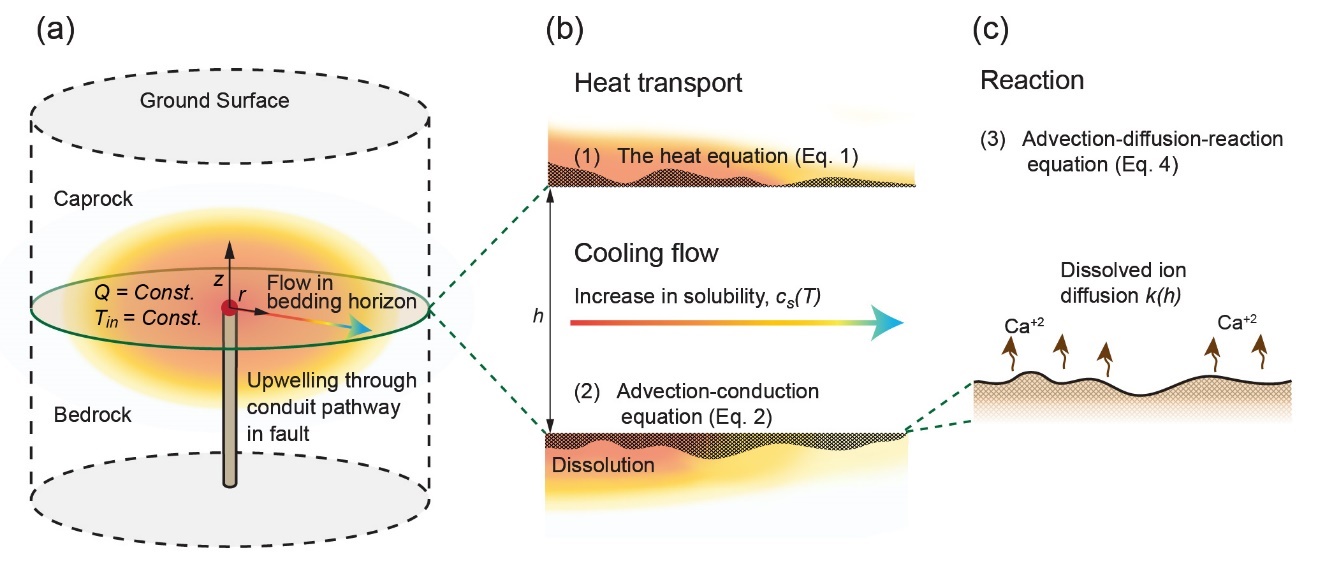


Fig. 2. **a** Geometry of numerical AHD model, **b** close-ups showing heat transport in the horizon, and **c** the surface reaction. **a** Fluid discharges from a pipe through an inlet of radius *r*0 (red point) with total flow rate *Q* and temperature *T*in. The flow is distributed radially, where it cools by transferring heat to the bedrock and caprock (color gradient arrow). **b** As the fluid cools, the saturation concentration *c*s(*T*) increases, inducing carbonate dissolution. Heat transport is governed by conduction in the rock and by the advection-conduction equation in the fluid, with heat exchange term *Θ* accounting for conduction between the rock and fluid. **c** Similarly, solute transport is calculated by using an advection-diffusion-reaction equation. For reaction, we assume a fully-transport-limited reaction that depends on the mass-transfer coefficient *k*(*h*).

We assume closed-system conditions and that the fluid pressure *p* in the confined aquifer exceeds the CO2 partial pressure PCO2 in the thermal water such that no degassing occurs. Under unconfined conditions, degassing of CO2 near the water table can diminish undersaturation and lead to supersaturation and precipitation21,45,46. Because cavities form as isolated voids, the overall permeability of the system essentially remains constant, allowing us to assume constant flow rate discharge27. Such an assumption breaks down only when dissolution approaches the surface outlet (i.e., when a breakthrough occurs)47.

Section S3 (SI) presents a theoretical analysis of dissolution in a uniform horizon as a result of heat exchange and with the rock and fluid cooling. The results of the analysis are then used to validate the numerical model presented here and implemented in a finite-difference code (see sections S4 and S2.3). Tables S1 and S4 give the notation for the physical parameters and their values, respectively.

* + 1. **Heat transport**

Heat transport in the rock, confining the horizon above and below (Fig. 2), is governed by the heat equation, which in polar coordinates takes the form

where *T* is the temperature, *t* is time, and *r* and *z* are the radial and vertical coordinates, respectively, both of which have their origin at the inlet. The quantity *αr* = *Kr/ρrCpr* is the thermal diffusivity, where the subscript *r* denotes rock, *K* is the thermal conductivity, *ρ* is the density, and *Cp* is the heat capacity48.

If heat transport along the horizon is governed by advection and conduction and complete mixing occurs along the horizon aperture, the “depth-averaged” heat transport equation can be written as

where *h* is the aperture, *q* is the fluid velocity integrated over *h* [L2/T] and calculated from the total volumetric flow rate *Q* using *q* = *Q*/(2*πr*), and *αf* is the fluid thermal diffusivity26,49. *Θ* accounts for the heat exchange with the bedrock and caprock and is calculated using Fourier’s law,

The complete mixing approximation can be validated *a posteriori* and is justified because the transverse temperature gradients in the horizon remain relatively small throughout speleogenesis.

* 1. **Reactive transport**

The depth-averaged solute transport advection-diffusion-reaction equation is

where *c* is the depth-averaged dissolved-calcite-ion concentration, *D* is the molecular diffusion coefficient, and *Ω* is the reaction term50,51. The transient terms in Eq. (4) are neglected, and the quasistatic approach is justified by the separation of timescales between mineral dissolution and flow and the relaxation of solute concentration (see ref. 52 and references therein).

Here, the initial CO2 partial pressure PCO2 is relatively high and the kinetics is rapid, so the rate-limiting step for the reaction is the diffusion of reaction products away from the mineral surface so that undersaturation is sustained26,53,54. The reaction term *Ω* is proportional to the difference between *c*s (*c* at calcite saturation or equilibrium at the given conditions) and *c*:

where *k*(*h*) is the mass-transfer coefficient, which is inversely proportional to the aperture *h*, *ε* is a geometrical correction term that accounts for the aperture inclination that develops following dissolution, and the factor of two accounts for the upper and lower surfaces. Section S2 (SI) provides the complete derivation of Eq. (5) and its justification.

The saturation concentration *c*s depends on the temperature and is expressed as a Taylor expansion to account for retrograde solubility:

where the constant *β* is evaluated from the data presented in section 4.1 and *T*0 is the reference (ambient) temperature. Finally, the change in the horizon aperture is calculated by using

where *ν* is a stoichiometric constant and *csol* is the soluble-solid concentration within the solid rock.

* 1. **Initial and boundary conditions**

The initial conditions are constant temperature, *T*0 = 20 °C. The boundary conditions at the horizon inlet (*r* = *r*0) are constant volumetric flow rate *Q* of fluid that has a temperature of *T*in = 60 °C and no initial undersaturation *θ* = 0, where *θ* = [*c*s(*T*)− *c*]*/*∆*c*s0and ∆*c*s0= *c*s(*T*0)− *c*s(*T*in). Here, we model conditions of a large domain, such that no substantial heat transport occurs near the boundaries and affects the results (*r* → *∞* and|*z*| → *∞*). In practice, the extent of *z* and *r* is limited to 1 km. A constant temperature *T*0 is set at the upper surface, and, at the horizon outlet (*r* = 1 km), the temperature and concentration are calculated by using the thermal and solute conservation equations (2) and (4) for a free-flow boundary. The remaining boundaries are assumed to be thermally insulating with zero conductive heat flux.

1. **Results and discussion**
   1. **Geochemical analysis**

Groundwater compositions encountered in deep-seated aquifers commonly demonstrate that PCO2 is approximately in equilibrium with the composition, temperature, and pressure conditions of the given rock. This equilibrium is promoted by the large CO2 fluxes originating, for example, from magmatic fluxes and the decomposition of organic matter33,55–57. The partial pressure PCO2 for pore water in sandstone and carbonate aquifers is typically much greater than that in crystalline aquifers33.

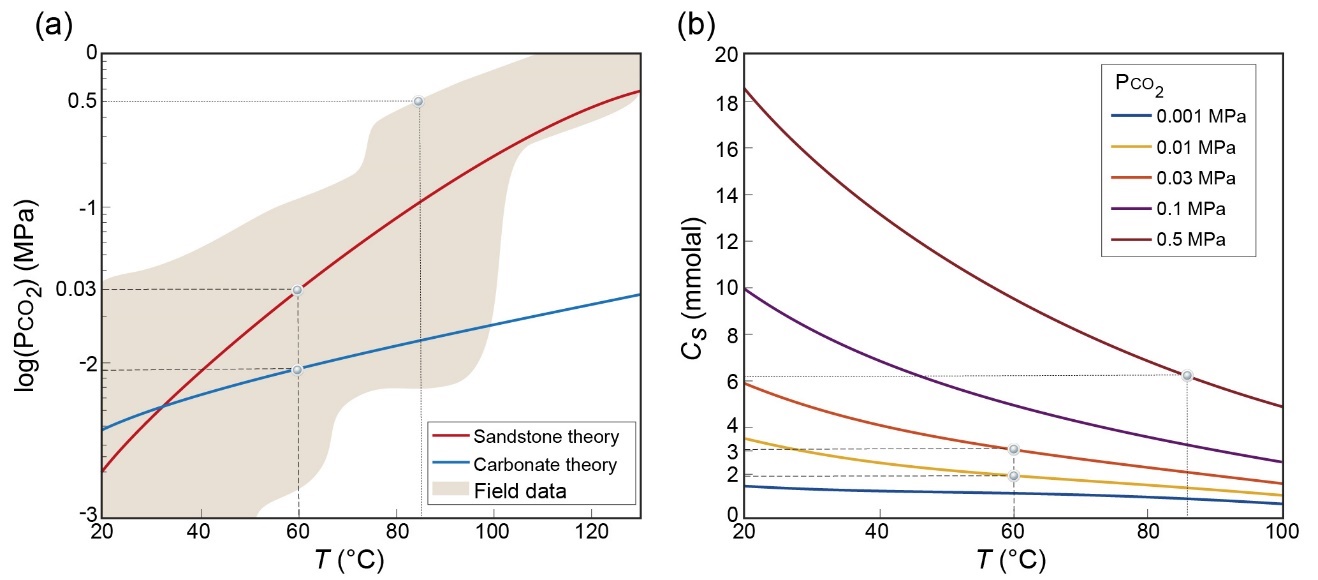
 Fig. 3. **a** Theoretical predictions of PCO2 as a function of temperature for clastic sandstone (red line) and carbonate (blue line) confined aquifers and the range of field data for sedimentary aquifers (carbonate and clastic rocks, brown region; modified from refs. 33,57). For pore water at *T* ≳ 30 °C, PCO2 progressively increases in sandstone aquifers above that in carbonate aquifers, with the difference reaching over 0.5 MPa for *T* ≳ 85 °C (see dotted lines). **b** Thermodynamic calculations of saturated calcite concentration as a function of temperature [*c*s(*T*)] and for different initial values of PCO2. Note that calcite retrograde solubility depends on initial PCO2: a greater initial PCO2 corresponds to a greater *c*s and the potential of the fluid to induce aggressive dissolution. Here, the inlet fluid temperature *T*in = 60 °C is considered in the numerical calculations and marked by dashed lines. Tables S2 and S3 (SI) present the chemical composition of aquifer water and of rock.

Figure 3a shows the theoretical predictions of PCO2 for sandstone (red line) and carbonate (blue line) aquifers obtained using PHREEKQC (v.3.7.0)58 and a range of field data for sedimentary aquifers (i.e., carbonate and clastic rocks, brown region)33,57. These calculations assume a confined aquifer and a closed system, and Tables S2 and S3 (SI) provide representative values of fluid pressure and composition. For slightly elevated temperatures (*T* ≳ 30 °C), PCO2 progressively increases for sandstone aquifers compared with carbonate aquifers.

Figure 3b plots the calcite concentration at saturation as a function of temperature for various initial values of PCO2, which reveals that calcite retrograde solubility depends strongly on PCO2. Thus, cooling water with high initial PCO2 leads to considerable calcite undersaturation, as indicated by the substantial increase in *c*s. Accordingly, for water from the sandstone aquifer at 60 °C (1.5–2 km depth) and with PCO2 ≈ 0.03 MPa that upwells and discharges into a shallow aquifer and cools to 20 °C, the saturation concentration can change by *c*s(*T* = 20 °C)− *c*s(*T* = 60 °C) ≈3 mmolal. In the carbonate aquifer under these conditions, PCO2 ≈ 0.01 MPa and *c*s(*T* = 20 °C) − *c*s(*T* = 60 °C) ≈1.5 mmolal (see dashed lines in Fig. 3b). For comparison, a typical value is *c*s = 2 mmolal59 in epigenic karst conditions under atmospheric pressure, demonstrating that these conditions favor karst formation.

Furthermore, field data show that, in sandstone pore water, PCO2 can be as high as 0.5 MPa for temperatures ≳85 °C, and *c*scan vary by over 12 mmolal when cooled to 20  °C (see dotted lines in Fig. 3). Under these conditions, a relatively small seepage of *Q* = 1 m3/day of cooling water dissolves 1 cubic meter of limestone rock of density *ρr* = 2400 kg/m3 within approximately 5 yr, which has the potential to form a cave within several hundred years. The upwelling of thermal groundwater from sandstone formations following long-range migration of the groundwater to shallower strata and carbonate aquifers is a common scenario60 that is likely relevant in our case study in Israel (see section 4.3). However, the long timescales of millions of years associated with hypogene karstification16,32 also permit speleogenesis with relatively low PCO2 (e.g., 0.001 MPa) and a small temperature decrease of only several degrees.

To summarize, this analysis indicates that carbonate dissolution due to cooling and retrograde solubility can be a major and sometimes highly aggressive karstification process that may form cave systems over relatively short timescales.

* 1. **Cave formation by localized dissolution**

The results obtained by solving the coupled equations (1)–(7) in the setting of Fig. 2 show that, upon flowing along the horizon, groundwater cools quickly (from 60 °C at the inlet to 25 °C at *r* = 10 m; Fig. 4a). The rapid cooling is promoted by the presence of a localized low-discharge source48, so the fluid reaches the quasi-steady state after about 50 yr. The fluid is hot and fully saturated at the inlet, *c* = *c*s(*T*in), increasing the undersaturation *θ* downstream and forming a characteristic convex shape (Fig. 4b). Undersaturation along the flow path is controlled by the interplay between (i) the reaction that reduces undersaturation, (ii) the progressive cooling that renews reactivity by elevating *c*s, and (iii) advection decaying as 1/*r* that removes reaction products and sustains undersaturation.

Near the inlet, the high advection and cooling rates cause a buildup of undersaturation. Further downstream, cooling and advection rates are much lower, and undersaturation decreases due to the dissolution reaction. Accordingly, the cave profile develops a pronounced maximum several meters from the inlet point (Fig. 4c). Over time, depending on the growth of the aperture *h*, the slowing rate of diffusion across the aperture also dampens the reaction. In turn, the slower reaction further increases the undersaturation, with the maximum undersaturation shifting progressively downstream (see arrow in Fig. 4b).

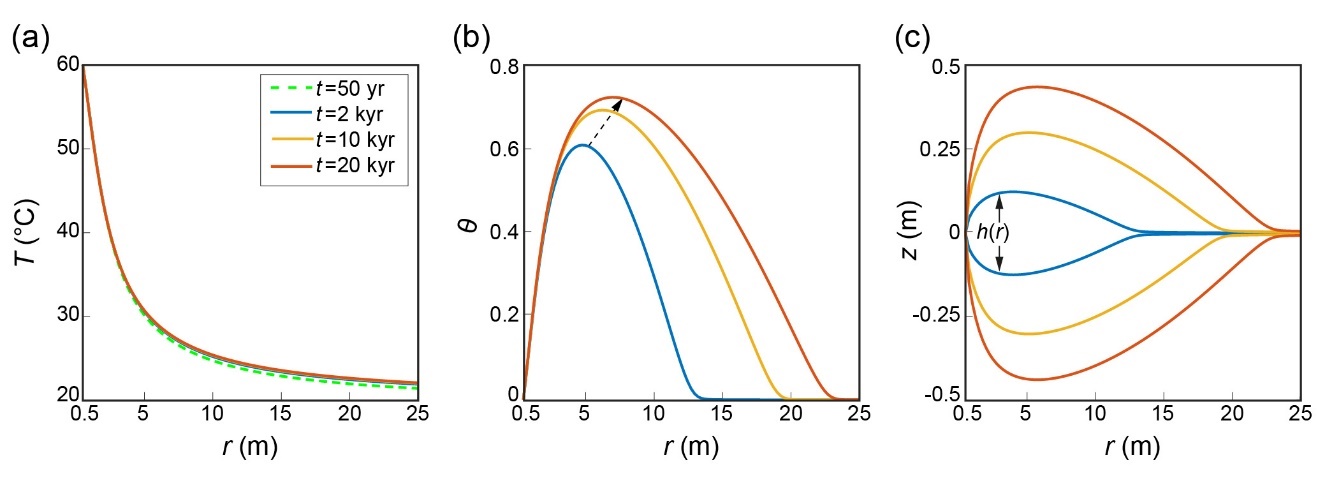


Fig. 4. Temperature *T*, solute undersaturation *θ*, and aperture profile *h* as functions of the radial position *r* along the horizon. **a** Upon flowing along the horizon, the water quickly cools, and the temperature curves overlap, revealing a relatively rapid (≈50 yr) approach to a quasi-steady thermal state. **b** Fluid cooling leads to a large increase in undersaturation, which first peaks and then decreases downstream because of the dissolution reaction. **c** Accordingly, the cave profile develops a pronounced maximum a few meters from the inlet. In addition, as the cave widens (i.e., *h* increases over time), the dissolution reaction becomes hindered by diffusion, resulting in a gradual increase in undersaturation, with its maximum shifting progressively downstream [see arrow in (**b**)].

The flow rate *Q* significantly affects karstification through the aspect ratio of the void, which is created because of dissolution. We define the degree *η* = *hp/lv* of dissolution focusing as the ratio between the maximum (peak) aperture *hp* and the inner length *lv* of the void calculated from the interval over which *h* > 1.01*h*0(Fig. 5). Lower flow rates elevate *η*, increasingly localizing the dissolution over time (measured here by the total fluid volume discharge, *Vf* = *Qt*). Higher flow rates result in a lower *η* (i.e., a more elongated void) due to a twofold advective effect: (i) heat is transported further downstream, leading to more gradual cooling and renewed reactivity, and (ii) the undersaturated fluid has a lower residence time and is flushed further down the flow path, thereby sustaining dissolution farther from the inlet.

The evolution of *η* shows a power-law dependence on the total volume *Vf* of fluid discharged and is inversely proportional to the flow rate, *η*(*Q*) = (*a/Q*)*Vfτ*, where *a* and *τ* are constants (see inset in Fig. 5). Our results emphasize that speleogenesis (high *η*) favors low flow rates; otherwise, only diffuse aperture variations or diagenetic alterations develop in the aquifer61–63. In addition, low flow rates are often associated with the upwelling of thermal flows from depth (>1 km) and thermal seepages and discharge from springs 17,60,64,65.

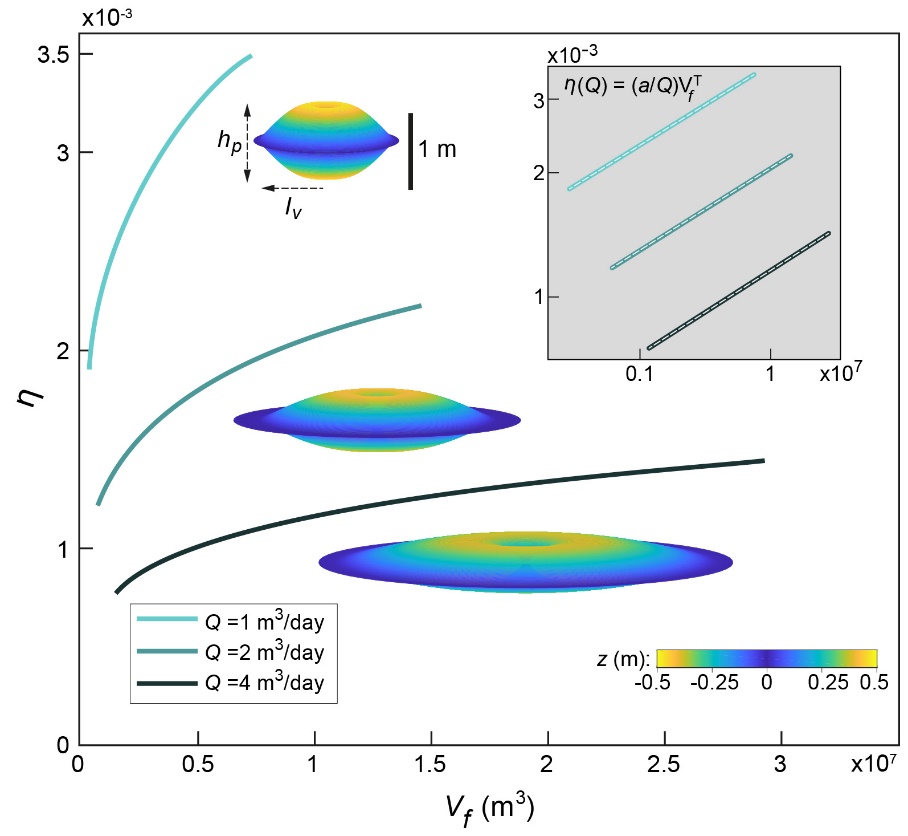


Fig. 5. Degree *η* of dissolution focusing as a function of fluid discharge volume *Vf* = *Qt* for different flow rates *Q*, between *t* = 1 to 20 kyr. *η* = *hp/lv*, where *hp* is the maximum aperture and *lv* is the inner length of the void calculated from the interval over which *h* > 1.01*h*0. Lower flow rates lead to a progressively larger *η* and localized dissolution, as demonstrated by the three-dimensional visualizations of the cavities after 20 kyr; the origin of the axis at the center and the vertical dimension *h* is rescaled by a factor of 30 relative to length. The log-log plots in the inset show the fits (gray dashed line) to *η*(*Q*) = (*a/Q*)*Vfτ* with *a* ranging from 1.45 to 1.8 × 10−4 and *τ* = 0.2.

The results indicate that the timescale of speleogenesis and the formation of human-size passages due to cooling of CO2-rich geothermal flow is several tens of thousands of years, which depends on the value assumed for PCO2. However, this indicates that the timescale of speleogenesis under these conditions is comparable to the typical epigenic speleogenesis timescales3,28 and can often be even shorter. This contradicts earlier estimates suggesting that hypogene speleogenesis requires exceedingly long timescales (e.g., ≈1 Ma)17,28. Specifically, early calculations of dissolution by rising thermal waters (ref. 28, p. 18) assumed rather low cooling rates and thermal gradients (5 °C/100 m). These led the author to conclude that “only under the most favorable conditions can dissolution by cooling of thermal water produce caves of traversable size. Even then, times on the order of 105 to 106 yr are required.”

* 1. **Confined speleogenesis and the case study**

The geometrical configuration of the caverns in the mathematical model is compatible with the field observations in our case study, which indicates that caves may form under confined conditions. Localized dissolution and speleogenesis require both a substantial geochemical driving force and an appropriate hydrogeological setup. The hydrogeological conditions that facilitate speleogenesis are (i) channelized upwelling flow through a fault by which the groundwater remains hot; (ii) aquifer confinement that promotes rapid cooling by a transition from channelized upwelling flow to lateral radial flow along the horizon and substantial heat exchange with the rock, as seen in Fig. 1; (iii) relatively low groundwater discharge that promotes localized dissolution and cave formation. Continuous low discharge is ensured by isolated caves, with no connection to the surface, so that the permeability of the overall system and the fluid flux remain essentially constant during speleogenesis27.

Conversely, the results indicate that, when thermal flow upwells through a fault and discharges to the surface (i.e., unconfined conditions), the advection rate increases with dissolution. This occurs either because of the overall system permeability and increased flow rate26 or because of localized dissolution channels that develop in zones of high advection rate49. In turn, high advection rates lead to low thermal gradients and promote more gradual dissolution along most of the fault or channels, and only toward the surface do the cooling and dissolution rates increase. However, in most instances under typical PCO2, it remains unclear whether a cave can form at shallow depth because of CO2 degassing and the concomitant solution supersaturation and precipitation that often occurs near the surface16,21. In fact, the upwelling of the flow that underlies the hypogenic processes inherently implies a certain degree of hydrogeological confinement because of the ubiquitous layered heterogeneity in sedimentary basins. Consequently, confined conditions are frequent for most hypogenic karstifications27.

Considering now the origin of upwelling flow and the geochemical driving force, note that the caves in this case study are assumed to form by upwelling from a clastic sandstone aquifer32,66. This aquifer is deeper (and hotter) than the carbonate aquifer and is shown herein to be associated with substantially higher PCO2 (section 4.1 and Fig. 3a). Nevertheless, the possible long time span of speleogenesis and the typically large timescale associated with tectonical and hydrogeological processes (on the order of 105 years and longer) prevents us from excluding a shallower carbonate aquifer as the origin of the upwelling groundwater. In particular, the large timescale permits the formation of the caves even by geothermal waters, which are only a few degrees warmer than the surrounding rock. Additionally, particularly large and spacious cave systems in the case-study group are presumably associated with large time spans or, alternatively, high-reactivity flow. These large voids can also become mechanically unstable have been associated with large documented collapses and sinkholes67,68.

* 1. **Cave morphology and pattern formation**

The analysis above indicates that cave formation due to confined thermal flows is feasible. We now investigate in more detail the characteristics of cave morphology. First, we present and discuss observations regarding the characteristic shape of caverns as predicted by the AHD model, with the maximum aperture located away from the feeder (Fig. 4c). This is followed by a brief analysis of the pattern and formation of cave passages. Smaller-scale morphologic features (meso-morphology) that are characteristic of hypogene processes (e.g., smooth walls, cupolas, and solution pockets) are also discussed27,32,43. However, we do not focus on such features because they often do not involve accurate and conclusive indicators of specific processes17,27,69.

* + 1. **Characteristic cave profile**

In the group of caves under investigation, cave feeders cannot be identified in most locations because they are covered by debris and sediment. Below, we present two unique observations of hypogene karst caves in Israel, in a setting like that of our case study (Fig. 6). Feeders are identified in these caves, allowing a direct comparison of the cave morphology with the results of section 4.2.

Cave (a) includes feeders and dissolution features along a prominent oblique fracture or fault (red dashed line in Fig. 6a), comprising the backbone of the three-dimensional maze cave66. While feeders and dissolution along the fracture appear clearly in many parts of the cave, the largest chambers and their ceiling peaks (A.1–A.5) are prominently located several meters away to 20 m away from the fracture and feeders (Fig. 6a). Similarly, observation in cave (b) reveals a characteristic convex profile with the most spacious region appearing approximately 10 m from the inlet, with diminishing passage sizes further downstream. See the plan-view and respective vertical profile a-a′ in Fig. 6b70.

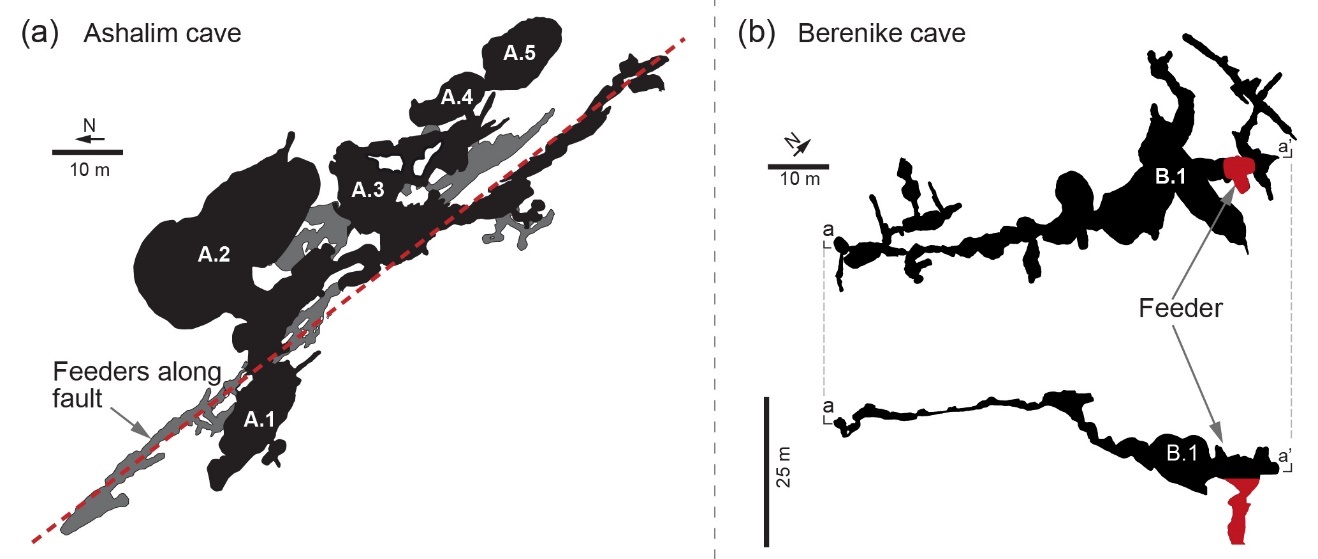


Fig. 6. Morphology of two hypogenic caves with identified feeders. **a** Feeders are visible in the Ashalim cave at the lower level along a fracture (dashed red line). The largest chambers (A.1–A.5) and their maximal ceiling height are prominently located several meters away to 20 m away from the fracture and feeders; see the gray zones marking overlapped lower levels66. **b** Plan view from above of upper level of the Berenike cave and its vertical profile a-a′ (below). The largest hall (B.1) is located approximately 10 m from the feeder (red region), and the cave has a convex shape with the passage size gradually decreasing downstream70.

Feeders are rare in caves from our case-study group, but the morphological characteristics associated with dissolution by thermal flow may be observed. However, these features are less clear and occur, for example, due to morphological reshaping by mechanical collapse and sedimentation71.

* + 1. **Maze-pattern formation**

Hypogene caves are often characterized by intricate maze or sponge-like patterns with many closed loops formed by a relatively uniform dissolution of the pre-existing void-space, which is typically composed of angular fracture networks11,17,27. This is in marked contrast with the common epigenic branchwork caves formed by competitive dynamics9,72. Several mechanisms for the genesis of maze caves have been suggested and studied for different hydrogeological setups, including epigenic and hypogenic origins (see the summary in refs. 17,28). These, among others, include uniform floodwater flow into fracture networks; for example, by a sinking stream73,74, or by distributed ascending groundwaters through the cave-forming layer, which enlarges the fractures uniformly and forms a maze pattern (“the transverse hypogenic cave origin model,” THCOM16,75–77).

The THCOM assumes a soluble rock layer sandwiched between permeable but insoluble strata and recharge from below. The insoluble strata sustain distributed flow and preclude the channelized flow that may develop in soluble rock due to preferential dissolution. The flow is then predominantly vertical and distributed, discharging via multiple feeders and enlarging all major fractures in the soluble rock at comparable rates. A prominent case-study example for the THCOM is the giant gypsum caves in Ukraine17,78. The THCOM setup requires a confined soluble layer between permeable and insoluble layers and cannot explain the formation of caves in aquifers confined by impermeable caprock layer, as in our case study. Nor can it explain the formation of caves in soluble carbonate rock successions, which remain largely obscure17,32. The incompatibility of the THCOM with our case study is also evident from the lack of multiple ceiling and floor feeders, which are invoked in the THCOM.

תמונה שמכילה טקסט

התיאור נוצר באופן אוטומטי

Fig. 7. Caves with characteristic maze-like patterns: **a** Chariton, **b** Qina32, and **c** results of a network model simulating dissolution by cooling and retrograde solubility in a fractured bedding horizon. Simulations produce a typical maze-like pattern characterized by numerous closed loops formed over 43 kyr. Only 1-m-wide passages or larger are shown. Figures S3 and S4 (SI) show the initial fracture distribution and the progression of speleogenesis up to the final state shown in (**c**).

Here, we present an additional model that extends the AHD model presented in section 3.1 to demonstrate maze-cave formation by radially cooling geothermal flow. This model simulates point injection of geothermal flow into a dissolving network of conduits in a fractured bedding horizon and considers cooling and retrograde solubility [see section S5 (SI) for details]. To run the model, we exploit the fact that the temperature distribution is essentially time-invariant, as demonstrated in section 4.1.

Compared with other speleological patterns such as common epigenic branchwork caves, the most prominent attribute of the maze pattern is the appearance of many closed loops17. This is seen in the plan views of the caves (Figs. 7a and b) and in the corresponding simulation results (Fig. 7c). The simulated maze pattern was obtained with a relatively low flow rate of *Q* = 2 m3/day, which is characteristic of thermal flows and seepages17,60,64,65. This result demonstrates the formation of maze caves without diffuse infiltration through an insoluble layer as commonly suggested17,27,28,77,79; it also recalls the inherent tendency of confined artesian flows to develop maze caves in carbonate rock successions.

An additional prominent result from the simulation is the merging of adjacent passages in some areas to form merged halls (bulky black regions, Fig 7c), which can explain the formation of some of the prevalent halls encountered in maze caves (see, e.g., Figs. 7a and b)32.

1. **Summary and Conclusions**

A geochemical analysis suggests that, because of a high initial CO2 partial pressure PCO2 and the retrograde solubility of carbonates, cooling of deep-seated geothermal flow can produce highly aggressive solutions and hypogene karstification. A numerical analysis further reveals that a channelized upwelling flow discharging to a confined aquifer promotes localized dissolution and speleogenesis. The localized dissolution is promoted by continuous low-flow-rate discharge, which is typical of upwelling thermal flows. In contrast with previous estimates, the present results show that such a process may be a common mechanism by which caves form on relatively short timescales of several tens of kyr.

Dissolution by this process further explains the location of caves and the major characteristics of the hypogene cave morphology observed in nature. The model predicts that the dissolution rate peaks a few meters downstream from the inlet at points corresponding to field observations. Moreover, the maze-pattern caves, which are typical of hypogene speleogenesis and have not been explained by previous models, are shown to form under confined conditions and in carbonate rock successions. Under these conditions, thermal water can flow from a vertical inlet into the network of conduits in a bedding horizon and create a maze cave through lateral flow that does not escape upwards.

From a broad perspective, this study emphasizes the link between global geothermal heat and the global carbon balances: Earth’s geothermal heat loss in conjunction with the large deep-seated CO2 fluxes induce transformations of the upper-crust carbonate strata and form extensive speleological systems, which in turn affect back CO2 fluxes.

**Acknowledgments**

We thank P. Lu for sharing the field data used in Fig. 3a.

**Author contributions**

R. Roded contributed by planning and conducting the research and writing the manuscript. E. Aharonov and P. Szymczak contributed by planning the research and writing the manuscript. A. Frumkin contributed by planning the research and guiding field trips. N. Weber and B. Lazar conducted the geochemical analysis and contributed to the writing of the manuscript.

**Competing interests**

The authors declare no competing interests

**Correspondence** and requests for materials should be addressed to E.A. or to R.R. (emails: [einatah@mail.huji.ac.il](mailto:einatah@mail.huji.ac.il); [roi.roded@mail.huji.ac.il](mailto:roi.roded@mail.huji.ac.il)).

**References**

1. Agar, S. M. & Geiger, S. Fundamental controls on fluid flow in carbonates: current workflows to emerging technologies. *Geol. Soc. Lond. Spec. Publ.* **406**, 1–59 (2015).
2. Jamtveit, B. & Hammer, Ø. Sculpting of rocks by reactive fluids. *Geochem. Perspect.* **1**, 341–342 (2012).
3. Dreybrodt, W. *Processes in karst systems, physics, chemistry, and geology*. (Springer, 1988).
4. Phillips, O. M. *Geological fluid dynamics: sub-surface flow and reactions*. (Cambridge University Press, 2009).
5. Aharonov, E., Whitehead, J. A., Kelemen, P. B. & Spiegelman, M. Channeling instability of upwelling melt in the mantle*. J. Geophys. Res. Solid Earth* **100**, 20433 (1995).
6. Aharonov, E., Spiegelman, M. & Kelemen, P. Three-dimensional flow and reaction in porous media: Implications for the Earth’s mantle and sedimentary basins. *J. Geophys. Res. Solid Earth* **102**, 14821–14833 (1997).
7. Aharonov, E., Tenthorey, E. & Scholz, C. H. Precipitation sealing and diagenesis: 2. Theoretical analysis. *J. Geophys. Res. Solid Earth* **103**, 23969 (1998).
8. Ortoleva, P., Chadam, J., Merino, E. & Sen, A. Geochemical self-organization II; the reactive-infiltration instability. *Am. J. Sci.* **287**, 1008–1040 (1987).
9. Szymczak, P. & Ladd, A. J. C. The initial stages of cave formation: Beyond the one-dimensional paradigm. *Earth Planet. Sci. Lett.* **301**, 424–432 (2011).
10. Bakalowicz, M. Karst and karst groundwater resources in the Mediterranean. *Environ. Earth Sci.* **74**, 5–14 (2015).
11. Ford, D. & Williams, P. *Karst hydrogeology and geomorphology* (John Wiley & Sons, 2007).
12. Huenges, E. & Ledru, P. *Geothermal energy systems: exploration, development, and utilization*. (John Wiley & Sons, 2011).
13. Fu, X., Cueto-Felgueroso, L. & Juanes, R. Pattern formation and coarsening dynamics in three-dimensional convective mixing in porous media. *Philos. Trans. R. Soc. Math. Phys. Eng. Sci.* **371**, 20120355 (2013).
14. Snæbjörnsdóttir, S. Ó. *et al.* Carbon dioxide storage through mineral carbonation. *Nat. Rev. Earth Environ.* **1**, 90–102 (2020).
15. Rattez, H., Disidoro, F., Sulem, J. & Veveakis, M. Influence of dissolution on long-term frictional properties of carbonate fault gouge. *Geomech. Energy Environ.* **26**, 100234 (2021).
16. Klimchouk, A. Types and settings of hypogene karst. In *Hypogene Karst Regions and Caves of the World* 1–39 (Springer, 2017).
17. Palmer, A. N. Distinction between epigenic and hypogenic maze caves. *Geomorphology* **134**, 9–22 (2011).
18. Audra, P. & Palmer, A. N. Research frontiers in speleogenesis. Dominant processes, hydrogeological conditions and resulting cave patterns*. Acta Carsologica* **44**, (2015).
19. De Vriendt, K., Pool, M. & Dentz, M. Heterogeneity‐Induced Mixing and Reaction Hot Spots Facilitate Karst Propagation in Coastal Aquifers. *Geophys. Res. Lett.* **47**, e2020GL087529 (2020).
20. De Waele, J. *et al.* Sulfuric acid speleogenesis (SAS) close to the water table: examples from southern France, Austria, and Sicily. *Geomorphology* **253**, 452–467 (2016).
21. Dublyansky, Y. Hydrothermal caves. In *Encyclopedia of caves* 546–552 (Elsevier, 2019).
22. Gabrovšek, F. & Dreybrodt, W. Early hypogenic carbonic acid speleogenesis in unconfined limestone aquifers by upwelling deep-seated waters with high CO2 concentration: a modelling approach*. Hydrol. Earth Syst. Sci.* **25**, 2895–2913 (2021).
23. Mylroie, J. E. Coastal caves. In *Encyclopedia of Caves* 301–307 (Elsevier, 2019).
24. Polyak, V. J., Hill, C. A., Asmerom, Y. & Decker, D. D. A Conceptual Model for Hypogene Speleogenesis in Grand Canyon, Arizona. In *Hypogene Karst Regions and Caves of the World* 555–564 (Springer, 2017).
25. Bakalowicz, M. J., Ford, D., Miller, T., Palmer, A. & Palmer, M. Thermal genesis of dissolution caves in the Black Hills, South Dakota. *Geol. Soc. Am. Bull.* **99**, 729–738 (1987).
26. Andre, B. J. & Rajaram, H. Dissolution of limestone fractures by cooling waters: Early development of hypogene karst systems. *Water Resour. Res.* **41**, (2005).
27. Klimchouk, A. Speleogenesis—Hypogenе. In *Encyclopedia of Caves* 974–988 (Elsevier, 2019).
28. Palmer, A. N. Origin and morphology of limestone caves. *Geol. Soc. Am. Bull.* **103**, 1–21 (1991).
29. Audra, P., Mocochain, L., Bigot, J.-Y. & Nobécourt, J.-C. Morphological indicators of speleogenesis: hypogenic speleogens. *Hypogene Speleogenesis Karst hydrogeology of artesian basins. pp 17À22. Simferopol Ukr. Inst. Speleol. Karstology Spec. Pap. 1* (2009).
30. Covington, M. D. & Perne, M. Consider a cylindrical cave: A physicist’s view of cave and karst science. *Acta Carsologica* **44**, (2016).
31. Kaufmann, G., Gabrovšek, F. & Romanov, D. Deep conduit flow in karst aquifers revisited. *Water Resour. Res.* **50**, 4821–4836 (2014).
32. Frumkin, A., Langford, B., Lisker, S. & Amrani, A. Hypogenic karst at the Arabian platform margins: Implications for far-field groundwater systems. *GSA Bull.* **129**, 1636–1659 (2017).
33. Coudrain-Ribstein, A., Gouze, P. & de Marsily, G. Temperature-carbon dioxide partial pressure trends in confined aquifers. *Chem. Geol.* **145**, 73–89 (1998).
34. Craw, D. Fluid flow at fault intersections in an active oblique collision zone, Southern Alps, New Zealand. *J. Geochem. Explor.* **69**, 523–526 (2000).
35. Hyman, J. D. & Dentz, M. Transport Upscaling under Flow Heterogeneity and Matrix-Diffusion in Three-Dimensional Discrete Fracture Networks. *Adv. Water Resour.* **155**, 103994 (2021).
36. Micklethwaite, S. & Cox, S. F. Progressive fault triggering and fluid flow in aftershock domains: Examples from mineralized Archaean fault systems. *Earth Planet. Sci. Lett.* **250**, 318–330 (2006).
37. Tripp, G. I. & Vearncombe, J. R. Fault/fracture density and mineralization: a contouring method for targeting in gold exploration. *J. Struct. Geol.* **26**, 1087–1108 (2004).
38. Ishibashi, T., McGuire, T. P., Watanabe, N., Tsuchiya, N. & Elsworth, D. Permeability evolution in carbonate fractures: Competing roles of confining stress and fluid pH. *Water Resour. Res.* **49**, 2828–2842 (2013).
39. López, D. L. & Smith, L. Fluid flow in fault zones: Influence of hydraulic anisotropy and heterogeneity on the fluid flow and heat transfer regime. *Water Resour. Res.* **32**, 3227–3235 (1996).
40. Auler, A. S. Caves and Speleogenesis in the Lagoa Santa Karst. In *Lagoa Santa Karst: Brazil’s Iconic Karst Region* 167–186 (Springer, 2020).
41. Filipponi, M., Jeannin, P.-Y. & Tacher, L. Evidence of inception horizons in karst conduit networks. *Geomorphology* **106**, 86–99 (2009).
42. Frumkin, A. & Gvirtzman, H. Cross-formational rising groundwater at an artesian karstic basin: the Ayalon Saline Anomaly, Israel. *J. Hydrol.* **318**, 316–333 (2006).
43. Frumkin, A. & Fischhendler, I. Morphometry and distribution of isolated caves as a guide for phreatic and confined paleohydrological conditions. *Geomorphology* **67**, 457–471 (2005).
44. Klimchouk, A. *et al.* Hypogenic origin, geologic controls and functional organization of a giant cave system in Precambrian carbonates, Brazil. *Geomorphology* **253**, 385–405 (2016).
45. Meakin, P. & Jamtveit, B. Geological pattern formation by growth and dissolution in aqueous systems. *Proc. R. Soc. Math. Phys. Eng. Sci.* **466**, 659–694 (2010).
46. Polyak, V., Hill, C. & Asmerom, Y. Age and evolution of the Grand Canyon revealed by U-Pb dating of water table-type speleothems. *Science* **319**, 1377–1380 (2008).
47. Szymczak, P. & Ladd, A. J. C. Wormhole formation in dissolving fractures. *J. Geophys. Res. Solid Earth* **114**, 1–22 (2009).
48. Stauffer, F., Bayer, P., Blum, P., Giraldo, N. M. & Kinzelbach, W. *Thermal use of shallow groundwater*. (CRC Press, 2019).
49. Chaudhuri, A., Rajaram, H. & Viswanathan, H. Early‐stage hypogene karstification in a mountain hydrologic system: A coupled thermohydrochemical model incorporating buoyant convection. *Water Resour. Res.* **49**, 5880–5899 (2013).
50. Hanna, R. B. & Rajaram, H. Influence of aperture variability on dissolutional growth of fissures in karst formations. *Water Resour. Res.* **34**, 2843–2853 (1998).
51. Szymczak, P. & Ladd, A. J. C. Reactive-infiltration instabilities in rocks. Fracture dissolution. *J. Fluid Mech.* **702**, 239–264 (2012).
52. Roded, R., Aharonov, E., Holtzman, R. & Szymczak, P. Reactive flow and homogenization in anisotropic media. *Water Resour. Res.* p.e2020WR027518. (2020).
53. Fu, X., Cueto-Felgueroso, L., Bolster, D. & Juanes, R. Rock dissolution patterns and geochemical shutdown of–brine–carbonate reactions during convective mixing in porous media. *J. Fluid Mech.* **764**, 296–315 (2015).
54. Golfier, F. *et al.* On the ability of a Darcy-scale model to capture wormhole formation during the dissolution of a porous medium. *J. Fluid Mech.* **457**, 213–254 (2002).
55. Ceriotti, G., Geloni, C., Dalla Rosa, M., Guadagnini, A. & Porta, G. Probabilistic modeling of field-scale CO2 generation by carbonate–clay reactions in sedimentary basins. *Hydrol. Earth Syst. Sci.* **25**, 3539–3553 (2021).
56. Cerón, J. C., Pulido-Bosch, A. & Bakalowicz, M. Application of principal components analysis to the study of CO2-rich thermomineral waters in the aquifer system of Alto Guadalentín (Spain). *Hydrol. Sci. J.* **44**, 929–942 (1999).
57. Lu, P., Luo, P., Zhang, G., Zhang, S. & Zhu, C. A mineral-water-gas interaction model of pCO2 as a function of temperature in sedimentary basins. *Chem. Geol.* **558**, 119868 (2020).
58. Parkhurst, D. L. & Appelo, C. *Description of input and examples for PHREEQC version 3: a computer program for speciation, batch-reaction, one-dimensional transport, and inverse geochemical calculations*. US Geological Survey (2013).
59. Dreybrodt, W., Gabrovšek, F. & Romanov, D. *Processes of a Speleogenesis: A Modeling Approach*. vol. 4 (Založba ZRC, 2005).
60. Garven, G. Continental-scale groundwater flow and geologic processes. *Annu. Rev. Earth Planet. Sci.* **23**, 89–117 (1995).
61. Jones, G. D. & Xiao, Y. Geothermal convection in the Tengiz carbonate platform, Kazakhstan: Reactive transport models of diagenesis and reservoir quality. *AAPG Bull.* **90**, 1251–1272 (2006).
62. Jones, G. D. & Xiao, Y. Geothermal convection in South Atlantic subsalt lacustrine carbonates: Developing diagenesis and reservoir quality predictive concepts with reactive transport models. *AAPG Bull.* **97**, 1249–1271 (2013).
63. Mazzullo, S. & Harris, P. Mesogenetic dissolution: its role in porosity development in carbonate reservoirs. *AAPG Bull.* **76**, 607–620 (1992).
64. Roded, R., Shalev, E. & Katoshevski, D. Basal heat-flow and hydrothermal regime at the Golan-Ajloun hydrological basins. *J. Hydrol.* **476**, 200–211 (2013).
65. Weber, N. *et al.* The circulation of the Dead Sea brine in the regional aquifer. *Earth Planet. Sci. Lett.* **493**, 242–261 (2018).
66. Frumkin, A. & Langford, B. Arid hypogene karst in a multi-aquifer system: hydrogeology and speleogenesis of Ashalim Cave, Negev Desert, Israel. *Geol. Soc. Lond. Spec. Publ.* **466**, 187–200 (2018).
67. Frumkin, A. *et al.* Gravitational deformations and fillings of aging caves: the example of Qesem karst system, Israel. *Geomorphology* **106**, 154–164 (2009).
68. Frumkin, A. *et al.* Sagging and collapse sinkholes over hypogenic hydrothermal karst in a carbonate terrain. *Geomorphology* **229**, 45–57 (2015).
69. Mylroie, J. E. & Mylroie, J. R. Diagnostic features of hypogenic karst: is confined flow necessary? In *NCKRI Symposium* vol. 1 12–26 (2009).
70. Ullman, M. Human selection and exploitation patterns of complex karstic caves during the later prehistory of the southern Levant. PhD in progress. (The Hebrew University, 2020).
71. Frumkin, A. *Holy Land Atlas: Judean Desert Caves*. (Magnes Press, 2015).
72. Worthington, S. R. H. & Ford, D. C. Self-organized permeability in carbonate aquifers. *Groundwater* **47**, 326–336 (2009).
73. Howard, A. D. & Groves, C. G. Early development of karst systems: 2. Turbulent flow. *Water Resour. Res.* **31**, 19–26 (1995).
74. Kaufmann, G. & Braun, J. Karst aquifer evolution in fractured rocks. *Water Resour. Res.* **35**, 3223–3238 (1999).
75. Klimchouk, A. Speleogenesis in gysum. *Int. J. Speleol.* **25**, 5 (1996).
76. Klimchouk, A. Conceptualisation of speleogenesis in multi-storey artesian systems: a model of transverse speleogenesis. *Int. J. Speleol.* **34**, 4 (2005).
77. Rehrl, C., Birk, S. & Klimchouk, A. Conduit evolution in deep-seated settings: Conceptual and numerical models based on field observations. *Water Resour. Res.* **44**, (2008).
78. Klimchouk, A. Large maze caves in gypsum in the Western Ukraine: speleogenesis under artesian conditions. *Natl. Speleol. Soc. Bull* **53**, 71–82 (1991).
79. Clemens, T., Hückinghaus, D., Sauter, M., Liedl, R. & Teutsch, G. Simulation of the evolution of maze caves. In *Proceedings, 12th International Congress of Speleology* vol. 2 65–68 (1997).