Geomorphology 204 (2014) 683–691

Contents lists available at ScienceDirect

Geomorphology

journal homepage: www.elsevier.com/locate/geomorph

'Looping caves' versus 'water table caves': The role of base-level changes and recharge variations in cave development

Franci Gabrovšek [a], Philipp Häuselmann [b,*,1], Philippe Audra [c,d]

[a] Karst Research Institute ZRC SAZU, Titov trg 2, 66230 Postojna, Slovenia
[b] Swiss Institute for Speleology and Karst Studies SISKA, c.p. 818, 2301 La Chaux-de-Fonds, Switzerland
[c] Polytech Nice-Sophia, University of Nice, Sophia Antipolis, 930 Route des Colles, 06903 Sophia-Antipolis, France
[d] Innovative City, IMREDD, Nice, France

ARTICLE INFO

Article history:
Received 22 February 2013
Received in revised form 17 September 2013
Accepted 21 September 2013
Available online 30 September 2013

Keywords:
Speleogenesis
Recharge variations
Looping cave
Water table cave
Modelling

ABSTRACT

The vertical organisation of karst conduit networks has been the focus of speleogenetic studies for more than a century. The four state model of Ford and Ewers (1978), which still is considered as the most general, relates the geometry of caves to the frequency of permeable fissures. The model suggests that the 'water table caves' are common in areas with high fissure frequency, which is often the case in natural settings. However, in Alpine karst systems, water table caves are more the exception than the rule. Alpine speleogenesis is influenced by high uplift, valley incision rates and irregular recharge. To study the potential role of these processes for speleogenesis in the dimensions of length and depth, we apply a simple mathematical model based on coupling of flow, dissolution and transport. We assume a master conduit draining the water to the spring at a base level. Incision of the valley triggers evolution of deeper flow pathways, which are initially in a proto-conduit state. The master conduit evolves into a canyon following the valley incision, while the deep pathways evolve towards maturity and tend to capture the water from the master conduits. Two outcomes are possible: a) deep pathways evolve fast enough to capture all the recharge, leaving the master conduit dry; or b) the canyon reaches the level of deep pathways before these evolve to maturity. We introduce the Loop-to-Canyon Ratio (LCR), which predicts which of the two outcomes is more likely to occur in certain settings. Our model is extended to account for transient flow conditions. In the case of an undulating master conduit, floodwater is stored in troughs after the flood retreat. This water seeps through sub-vertical fractures ('soutirages') connecting the master conduit with the deep pathways. Therefore, the loops evolve also during the dry season, and the LCR is considerably increased. Although the model is based on several approximations, it leads to some important conclusions for vertical organisation of karst conduit networks and stresses the importance of base-level changes and transient recharge conditions. It therefore gives an explanation of speleogenesis that relies much more on the dynamic nature of water flow than on the static fracture density.

1. Introduction

For most of the last century, the question whether caves originate above, at, or below the regional karst water table was unsolved (see detailed discussion in Gabrovšek, 2000; Dreybrodt et al., 2005; Audra and Palmer, 2013). Field evidence and calculations did not give priority to any one of the conceptual models. Martel (1921) argued that cave enlargement is most intensive in the vadose zone where the infiltrating water is still aggressive, and because the high flow velocity enhances erosion. Davis (1930) and Bretz (1942) concluded from field studies that caves developed along Darcy flow paths below the water table. Swinnerton (1932) contended that caves are more likely to form where the density of groundwater flow is highest, that is, at and just below the water table (Fig. 1). Therefore, he considered the zone of water-table fluctuation to be the most favourable for cave origin. All the theories were partially supported by field observations.

The four state model (Ford, 1971; Ford and Ewers, 1978) relates fissure frequency and cave pattern in the dimension of length and depth (Fig. 1): Deep phreatic loops evolve when fissure frequency is low (e.g. few and widely spaced penetrable fissures), and 'water table caves' originate in cases of high frequency of permeable fissures. All the intermediate cases are also included in the model. The model was later expanded to six states (Ford and Williams, 1989), with one end-member being an isotropic rock with no fissures and thus no caves (for example a well-recrystallized marble) and the other one an isotropic highly porous rock where the porosity is so large that no distinct caves are formed (for example chalk). The four state model therefore answered

* Corresponding author.
E-mail addresses: gabrovsek@zrc-sazu.si (F. Gabrovšek), praezis@speleo.ch (P. Häuselmann), audra@unice.fr (P. Audra).
[1] Tel.: +41 77 426 2390.

0169-555X/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.geomorph.2013.09.016

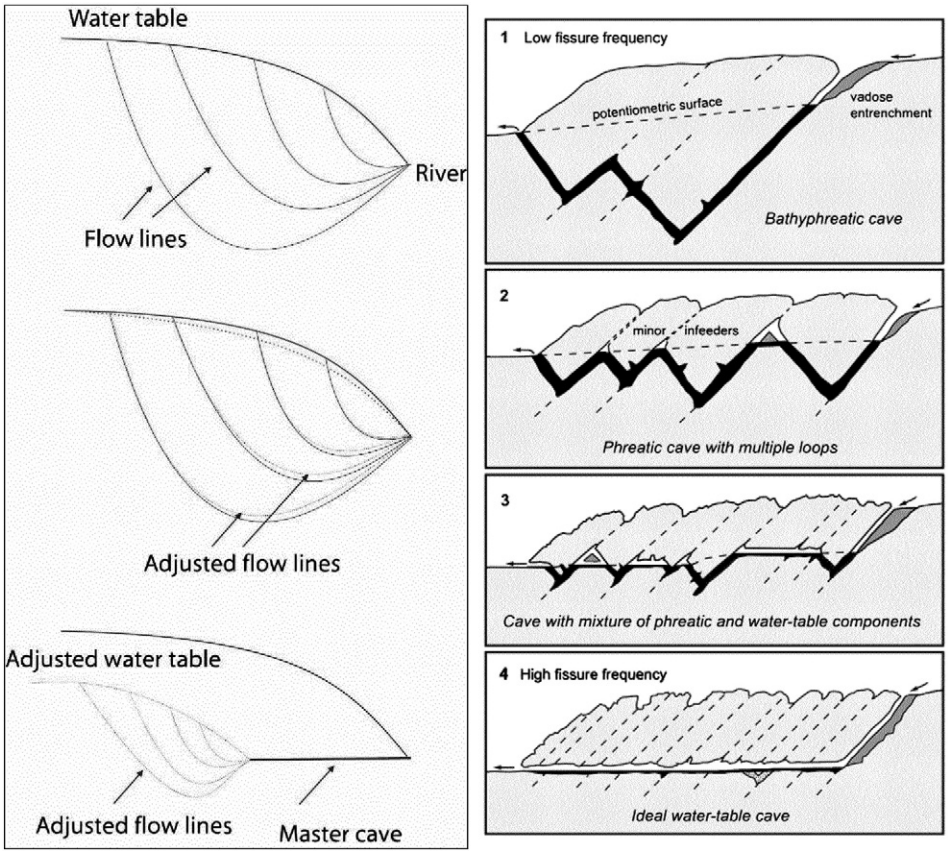

Fig. 1. Left: the water table cave hypothesis proposed by Swinnerton (1932). Right: The four state model of Ford and Ewers (1978). See text for discussion.

the main question in an easily understandable and quantifiable way. It therefore quickly became the main speleogenetic model, and is still regarded valid by many researchers: "The argument concerning whether caves formed above, at, or below the water table that so preoccupied researchers of the classic period, was definitely put to rest by Ford and Ewers (1978). The answer was, "yes"" (White, 2000).

However, continuing research on speleogenesis revealed several questions that were not compatible with the four state model. Worthington (2004, 2005) questioned the validity of the Ford–Ewers model by noting the development of sub-horizontal caves as much as 100 m below the water table. He also showed statistically that the depth of phreatic cave development is proportional to the overall length of flow paths and angle of the stratal dip.

The main point is that tectonised and fractured Alpine rocks should show many more water table and near water table caves than, for example, the relatively undisturbed limestones of the Mammoth Cave Plateau (USA). But in the Alps, there are very few water table caves or caves of State 3 (see Fig. 1 for an illustration). For example, the folded and thrusted Vercors massif (French Prealps) contains 277 large caves (from Lismonde and Frachet, 1978, modified), of which 190 are vadose and 15 unrecognizable (speleothems, breakdowns, etc.). The other 72 caves of phreatic origin are composed of 6 water table caves (8%) and 66 'looping caves' (92%). Another example is Hölloch cave (Switzerland), which by Ford and Ewers (1978) is considered as the type locality for State 2, although it is located in steeply dipping and densely fractured limestone. This topic was discussed in some length in Jeannin et al. (2000), where Derek Ford states in a public comment after the article: "The Four State Model … does not attempt to predict what will be the effective fissure frequency and aperture in any particular topographic or geologic setting …". In other words, the four state model cannot be used as a predictive model explaining "why did that particular cave form in this state at that location".

Ford and Ewers (1978) stipulate that fissure frequency increases with geological time, causing multilevel caves to evolve from State 1 to near State 4. However, even if the increase in fissure frequency over the lifespan of a karst seems reasonable enough (erosional unloading of the surface, tectonic release when valleys are deepened, continuous karstification of pre-existing small fissures, etc.) such effects are usually localized near the surface (creating the 'epikarst') or constrained to well-defined fissures. Thus, the implied time-dependency of the four state model induced confusion among many karst researchers who had troubles matching the model to the observed reality.

Palmer (1991) suggested that the plan pattern of caves is also controlled by discharge fluctuations, a view that is now widely accepted and also cited by Ford (1999). However, Ford (1999) does not take into account possible recharge variations while explaining the cave pattern in length and depth. During the last two decades, it has been found that recharge variations have a huge influence on speleogenesis. Water chemistry measurements as well as direct observations in caves have shown that floodwaters are much more corrosive and erosive than low waters (Palmer, 2000). Scallop size depends on flow velocity, and the velocity back-calculated from scallops also reveals that erosion mainly occurs during flood events (Lauritzen et al., 1983). Audra (1994) emphasized the influence of the epiphreatic (floodwater) zone for speleogenesis of passages of apparent phreatic origin, and Häuselmann et al. (2003) subsequently refined the model, explaining the speleogenesis of Bärenschacht (Switzerland) on the basis of floodwater fluctuations. We thus can ascertain that floodwater effects are very important in speleogenesis.

Furthermore, karst in orogens is subjected to rapid base-level changes which result in time-varying boundary conditions for the development of karst networks. The existence of cave levels has been related to stillstands of base level, however we still lack some basic understanding of how karst systems adopt to changes of erosional base level.

The first modelling studies of conduit network development in the dimensions of length and depth were presented by Gabrovšek and Dreybrodt (2001) and later by Kaufmann (2003) and Dreybrodt et al. (2005). They modelled the evolution of conduit networks in an unconfined fractured aquifer with constant recharge and/or constant head conditions. Flow and dissolution focuses at the position of the water table creating a highly permeable fringe which drops in time towards the position of a base level. The preferential outcome of all these models is a water table cave at the base level developing from the spring towards the interior of a massif. Deep phreatic loops resulted only when a pathway of prominent fractures was subjected to constant head conditions, or when a cycle of down-cutting and backfilling of the base level was introduced in the model. These results therefore again raise the question why the water table caves are not more abundant.

Here we focus on two factors which may control the evolution of karst networks in the dimension of length and depth, the role of base-level incision, and a scenario of transient recharge.

2. Hypothesis

In this section we present the initial hypotheses which have led to this work. These are based on field studies in Alpine systems and primarily concern the role of transient recharge in the vertical development of karst networks. Fig. 2 shows a system with a slightly undulating passage formed in the phreatic zone along the most permeable fracture pathways. After base-level incision, the passage has remained above the valley. Below the passage, a relatively undeveloped system of fissures and prominent fractures is assumed. The left column shows the evolution of the system driven by large recharge variations, with frequent flooding of the passage and dry periods in between. Flood water is shown in dark blue and the water that is left in the system after the flood in cyan. The right column shows a scenario with a uniform recharge. Fig. 2a–d presents the evolution of system in time for both scenarios. Note that the columns are not synchronous, they merely present the sequence of anticipated events. Fig. 2a shows an initial situation after the uplift and/or down-cutting. To the left, the water stays in the troughs of the main passage after flood recession; to the right continuous flow along the main passage is sustained. In both cases a steep gradient along yet undeveloped pathways connecting the main passage to the springs at the base level builds up, as denoted by arrows. Due to dissolutional widening, these pathways capture flow and empty the trapped flood water with increasing efficiency. The term 'soutirage' has been used for such passages (Häuselmann et al., 2003).

In the transient-recharge scenario (Fig. 2a, b, left), the main passage evolves only during floods, however soutirages are evolving also during the dry period, as long as the water remains stored in the troughs. In a constant recharge scenario (Fig. 2b, c, right), the main passage and soutirages evolve continuously. In Fig. 2c, the soutirages closest to the valley have developed enough to capture and direct all available flow to the lower spring. The process of flow capturing by soutirages evolves

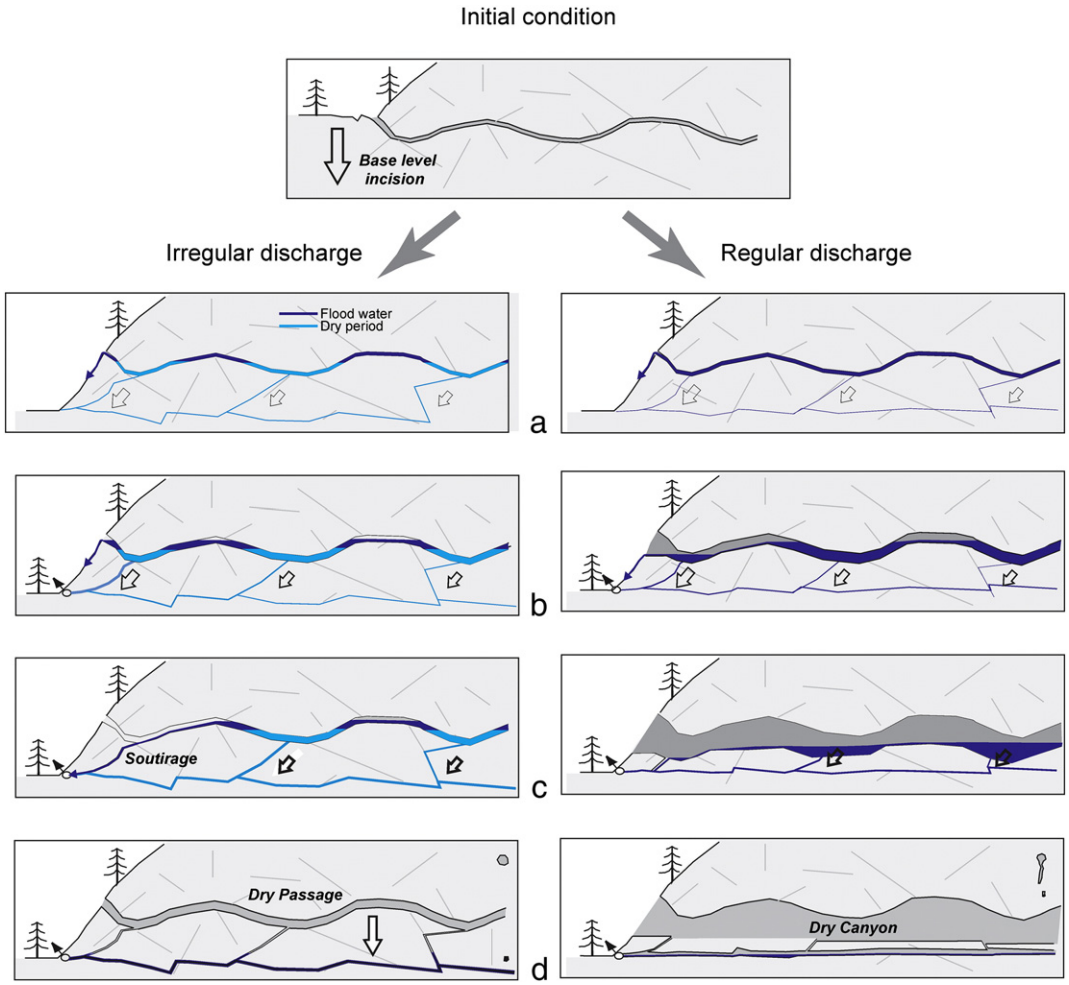

Fig. 2. The initial hypotheses stressing the importance of the recharge variations. Left and right columns show the evolution of a karst drainage network with irregular and regular recharge, respectively, at different stages of evolution after a quick incision of a base level (see text for discussion).

upstream. Finally, as shown on the left side of Fig. 2d, the hypothetical end-member of the floodwater cave is created: the lower passage has enlarged sufficiently to accommodate all the floodwater. This is equilibrium sensu Worthington (1991), attained (due to the rarity of high floods) only asymptotically and theoretically — since usually valley incision is much faster. On the right, the pathways connecting a passage to the lower spring have developed sufficiently to drain all available recharge. As usually, the fracture closest to the spring was enlarged first, capturing the flow completely. Canyon incision still continues upstream, until the next upstream pathway is sufficiently enlarged to drain all available flow. The cave therefore shows abandoned canyons that get deeper towards the interior, while a lower water table cave of partially phreatic origin is formed at the elevation of spring.

High recharge variations occur due to sporadic input (snowmelt, storm events), allogenic recharge from a surface catchment, low water retention in the uppermost layer, etc. High recharge variations have to be considered typical for all mountain belts (orographic rainfall, low infiltration in the winter time and intense snow melt in the spring–summer; as well as the folded/faulted structure which juxtaposes non-carbonate to carbonate rocks, favouring allogenic inputs) as well as for some tropical and Mediterranean settings. Low recharge variations are more typical in areas with thick soil cover tempering the effect of heavy rains, with porous non-carbonate rocks above the aquifer or in areas with very uniform climate. These conditions are much more rarely encountered; the evoked 'uniform climate' with constant daily rain, for instance, possibly does not exist anywhere on Earth. It follows that water table caves should be very rare in Alpine environments, and not very often encountered in other settings. An overview of cave maps of Europe, made by the authors, reveals that this is the case. But since we are dealing with geological time spans where the effects of changing climate and assessment of valley deepening rates are difficult, if not impossible, to constrain, we apply a simple mathematical model of conduit evolution in a karst system. To test the reasoning above, we built a simple mathematical model which assesses the relative importance of valley deepening rates, dissolution rates in karst, and a recharge scenario for the resulting structure of the karst aquifer in the dimensions of length and depth.

3. A conceptual model: Development of a karst network in areas of continuous deepening of the base level

To confirm and extend the hypotheses given above, we introduce a simple conceptual model with several parameters related to the base-level down-cutting and to the evolution of a karst network.

Fig. 3a–c shows a simplified cross-section of a karst massif at three different stages. Initially (Fig. 3a) a major conduit C1 drains water from the massif to the valley on the left. The conduit C1 and the initial position of the valley are on the Level 1. Below C1, deeper flow pathways, following set of fractures are possible. One of them is the pathway C2 which mainly extends along Level 2, at an initial depth Δz below the Level 1, and is connected to the recharge point on one side (point A) and to the spring in the valley. It is initially a proto-conduit with a laminar flow. Its evolution is driven by an aggressive solution entering along C1 at point A and by the hydraulic head, determined by the difference in the elevations between the bottom of C1 at point A and the position of the base level (Fig. 3b). Dotted grey lines mark other possible fractures along which several competing pathways may evolve. Here we analyse the development of one pathway (C2), which is either the only connecting inputs to outputs or being most permeable.

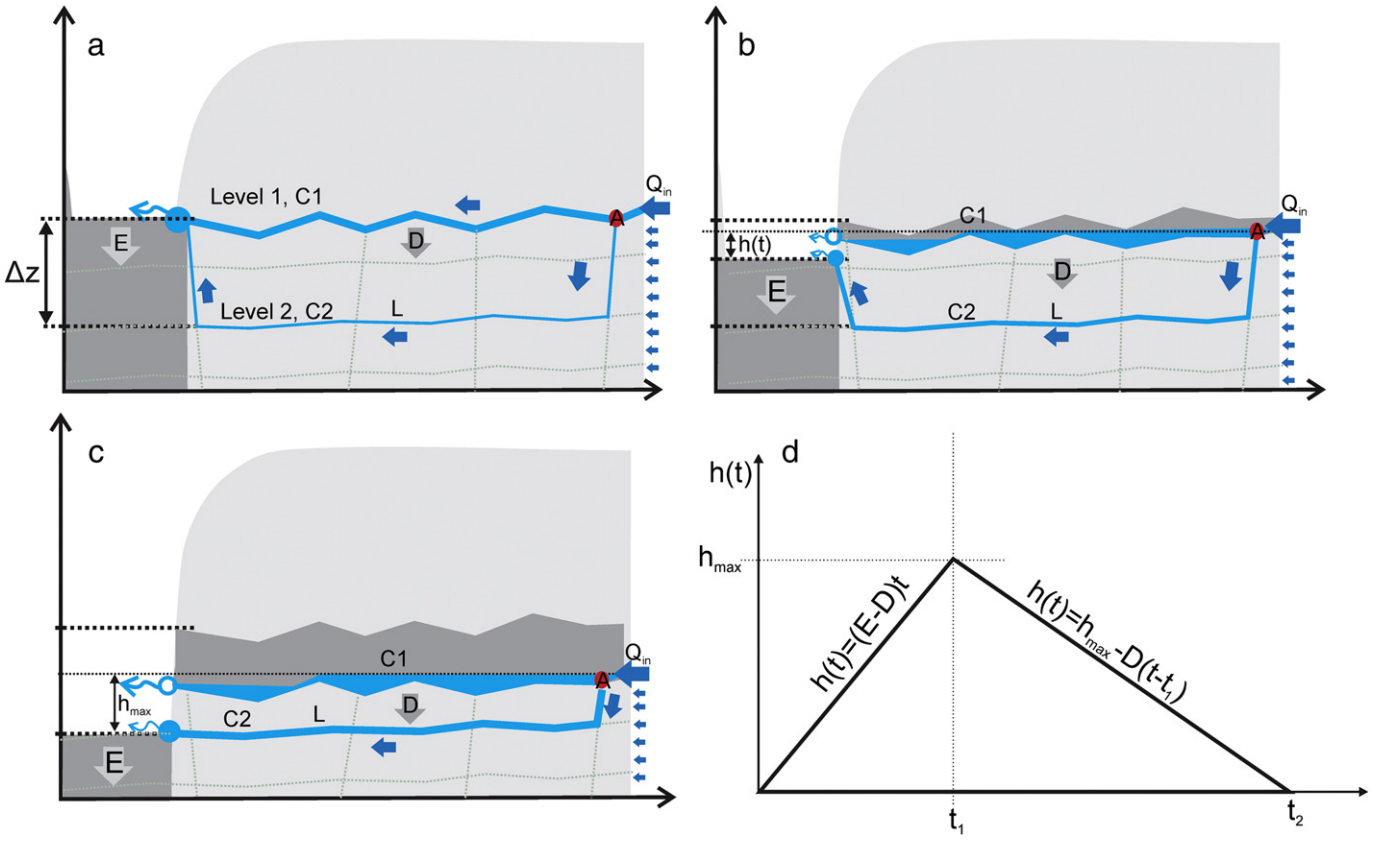

Fig. 3. Cross-section of an evolving karst massif with a base valley entrenching on the left. a) Initial state. b) Intermediate state. c) State with a maximal head difference along C2. d) Evolution of the hydraulic head acting on conduit C2. Description and role of all parameters are given in the text and Table 1.

Initially (Fig. 3a), there is practically no hydraulic head acting on C2, so it evolves extremely slowly or not at all. We assume that the valley deepens at some constant rate E [LT^{-1}], while C1 incises at a rate D [LT^{-1}]. If the incision of C1 is solely dissolutional, it is defined by the saturation state of the inflowing solution at point A. However, the incision of C1 can also be enhanced by a mechanical erosion of the sediment particles in water. C1 evolves into a canyon (see sequence in Fig. 3). E and D are constant in time; to assure flow out of the massif (i.e. from right to left in Fig. 3), the condition $E > D$ is assumed.

Deepening of the valley triggers the hydraulic head along C2. Its time dependence is shown in Fig. 3d and is given by:

$$h(t) = \begin{cases} (E-D)\cdot t = h_{\max}\dfrac{t}{t_1}; t \le t_1 \\ h_{\max} - D\cdot(t-t_1) = h_{\max}\dfrac{t_2-t}{t_2-t_1} t_1 < t < t_2 \end{cases}, \quad (1)$$

where

$$t_1 = \frac{\Delta z}{E}, t_2 = \frac{\Delta z}{D}, h_{\max} = (E-D)\cdot t_1.$$

Two possible outcomes of given scenario can be foreseen:

1. Canyon: C1 incises continuously until it becomes a canyon filling the space between Level 1 and Level 2.
2. Loop: C2 evolves fast enough to capture all the inflow at point A before C1 reaches the Level 2. C1 stops growing and remains a separate channel.

From here on we apply and derive some of the basic modelling equations, which are required to obtain a process-based assessment of the dilemma given above. To make the text self-contained, we first review some basic concepts and results of a single conduit evolution.

4. Modelling the evolution of a single karst conduit under variable hydraulic head difference

4.1. Dissolution rates along a uniform fracture

Groundwater flowing through fractures and conduits in soluble rocks dissolves and erodes their walls. We present a mathematical description of conduit evolution by coupling dissolution rate laws to the governing and conservating laws of flow and transport. Dissolution rates of karst rocks are generally a function of water saturation state with respect to the rock-forming minerals (Bögli, 1980; Dreybrodt, 1988). In limestone, dissolution rates drop linearly with increasing saturation ratio until a critical concentration at about 90% of saturation is reached (Eisenlohr et al., 1999; Kaufmann and Dreybrodt, 2007). From there onwards, rates follow a non-linear rate law with rate order $n = 4$–11, depending on the rock (Eisenlohr et al., 1999).

Coupling flow, dissolution and transport results in an exponential decrease of dissolution rates along a pathway (i.e. a single fracture or a succession of fractures) for a linear rate law, and a hyperbolic decrease for a non-linear rate law (Dreybrodt, 1996; Dreybrodt and Gabrovsek, 2002; Dreybrodt et al., 2005):

$$F(x) = \begin{cases} F_{in}\exp(-x/\lambda_1) \text{ for } x < x_s, \text{where } \lambda_1 = Qc_{eq}/Pk_1 \\ F(x) = F_{in}[1 + x/\lambda_n]^{\frac{n}{1-n}} \text{for} x_s \le x \le L, \text{where} \lambda_n = \dfrac{Qc_{eq}\left(1-c_{in}/c_{eq}\right)^{1-n}}{Pk_n(n-1)} \end{cases}. \quad (2)$$

The dissolution rate $F(x)$, where x denotes the flow distance along the pathway, is given in N / ($L^2 \cdot T$). F_{in}[N / ($L^2 \cdot T$)] is a dissolution rate at the entrance of a fracture. Parameters λ_1 and λ_n are the length scales for linear and nonlinear dissolution rate laws. The term 'penetration length' is used for them from here on. Q[L^3/T] is the flow rate, c_{in} and c_{eq} [$N \cdot L^{-3}$] are the initial and equilibrium concentration of calcium in the water, k_1 and k_n the linear and nonlinear rate constants and P the fracture perimeter. Widening rate of a fracture at distance x from the entrance is given by $da(x,t) / da = 2\gamma F(x,t)$, where $a(x,t)$ is the aperture width and γ converts the dissolution rate in mol/($cm^2 \cdot s$) into the rate of wall retreat in cm/y. For the meaning and typical values of all parameters and constants in this work, refer to Table 1.

When a pathway evolves under a constant head difference, the increase of flow rate increases the penetration length, which in turn causes an even faster increase of flow and dissolution rates. Such a feedback loop ends in an abrupt increase of both, called 'the breakthrough'. After the breakthrough, flow becomes turbulent and dissolution rates become increasingly uniform along the whole pathway (i.e. penetration lengths become large compared to the conduit length). Further increase of flow rates has a small effect on the dissolution rates, therefore $F(L) \approx F_{in}$. Fig. 4 presents the evolution of flow rate in time for a single wide fracture with initial aperture width of 0.02 cm, length of 1 km and the constant head difference of 50 m. The right vertical axis shows the cumulative volume of flow through the fracture presented by the grey dashed line.

The flow rate after the breakthrough increases for several orders of magnitude in a short time. In our scenario (Fig. 3), we expect that the deep pathway C2 will capture all the available recharge if it breaks through before C1 incises to Level 2. Here we follow the work of Dreybrodt and Gabrovšek (2000), who derived analytical approximation for the breakthrough time of a single fracture under constant head conditions, and extend the solution for the case of time varying hydraulic head as given in Eq. (1).

4.2. Analytical approximation of a fracture evolution under variable hydraulic head

For most natural scenarios, the following approximations are valid almost until the breakthrough: $\lambda_1 \approx 0$, $x_s \approx 0$ and $L/\lambda_n \gg 1$.

Table 1
Parameters and constants used in this work, their notation and units. Typical values are given for those which are constant in all presented cases.

Name of parameter	Notation and units	Typical values (if used)
Incision rate of the channel C1	D [$L \cdot T^{-1}$]	
Rate of base level lowering	E [$L \cdot T^{-1}$]	
Initial elevation difference between Level 1 and Level 2	Δz [L]	
Flow rate	Q [$L^3 \cdot T^{-1}$]	
Fracture perimeter	P [L]	
Length of a fracture/pathway	L [L]	
Hydraulic head	h [L]	
Initial aperture width	a_0 [L]	Few tenths of a millimetre
Linear rate constant	k_1 [$N \cdot L^{-2} \cdot T^{-1}$]	4×10^{-11} mol/cm^2/s
Rate order, Nonlinear rate constant	n, k_n [$N \cdot L^{-2} \cdot T^{-1}$]	4.4×10^{-11} mol/cm^2/s
Equilibrium concentration	c_{eq} [$N \cdot L^{-3}$]	10^{-6} mol/cm^3–2×10^{-6} mol/cm^3
Input concentration, Initial rates	c_{in} [$N \cdot L^{-3}$], F_{in} [$N \cdot L^{-2} \cdot T^{-1}$]	
Density, viscosity, gravitational acceleration		$\rho = 1$ g/cm^3, $\eta = 0.012$ cm/s, $g = 981$ cm^2/s

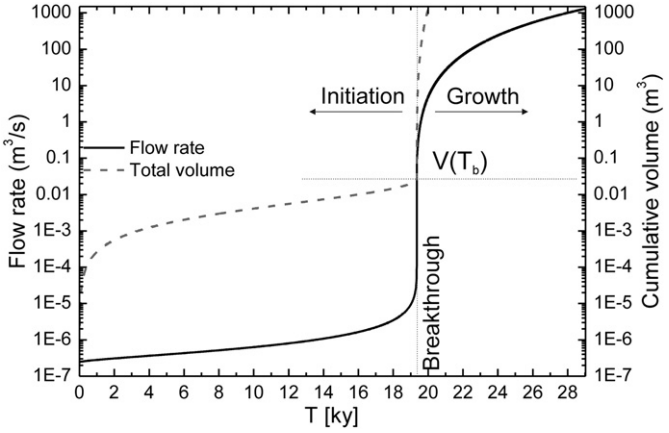

Fig. 4. Evolution of flow rate and cumulative volume of flow in a fracture with initial aperture width $a_0 = 0.02$ cm, length $L = 1$ km, hydraulic head $h = 50$ m. Black solid line presents temporal flow rate and dashed grey line the total volume of water that has passed the fracture since the beginning of its evolution.

Furthermore, almost the entire drop in dissolution rate occurs close to the entrance, so that most of a pathway is widened at a rate not much higher than that at the exit. Dreybrodt (1996) and Dreybrodt and Gabrovšek (2000, 2002) applied these approximations to derive an analytical result for the evolution of a uniform wide fracture. Taking a nonlinear rate law along the entire fracture and $F(x) = F(L)$ for $0 \le x \le L$, a differential equation describing the widening in time is obtained:

$$\frac{da}{dt} = 2\gamma F(L,t) = 2\gamma F_{in}\left(\frac{L}{\lambda_n(x_s t)}\right)^{\frac{n}{1-n}} \tag{3}$$

where

$$\lambda_n(t) = \frac{\rho g}{24\eta}\frac{a(t)^3 h(t)}{L}\frac{c_{eq}\left(1-c_{in}/c_{eq}\right)^{1-n}}{k_n(n-1)}. \tag{4}$$

λ_n contains time-dependent hydraulic head and aperture width. See Table 1 for the meaning and typical values of all parameters and constants. Introducing the penetration lengths for an initial aperture $a = a_0$ and maximal head $h = h_{\max}$

$$\lambda_0 = \frac{\rho g}{24\eta}\frac{a_0^3 h_{\max}}{L}\frac{c_{eq}\left(1-c_{in}/c_{eq}\right)^{1-n}}{k_n(n-1)}, \tag{5}$$

Eq. (4) is rewritten as:

$$\lambda_n(t) = \lambda_0\left(\frac{a(t)}{a_0}\right)^3\frac{h(t)}{h_{\max}}, \tag{6}$$

and the differential Eq. (3) becomes:

$$\frac{da}{dt} = 2\gamma F_{in}\left(\frac{L}{\lambda_0}\right)^{\frac{n}{1-n}}\left(\frac{a(t)}{a_0}\right)^{\frac{3n}{n-1}}\left(\frac{h(t)}{h_{\max}}\right)^{\frac{n}{n-1}}. \tag{7}$$

The differential equation is separable. Integration over a and some algebraic rearrangement gives explicit time dependence of the aperture width a:

$$a(t) = a_0\left[1-2\cdot\gamma\cdot F_{in}\left(\frac{L}{\lambda_0}\right)^{\frac{n}{1-n}}\cdot\frac{2n+1}{n-1}\cdot a_0^{-1}\cdot h_{\max}^{\frac{-n}{n-1}}\int_0^t h(\tau)^{\frac{n}{n-1}}d\tau\right]^{\frac{1-n}{2n+1}}. \tag{8}$$

$a(t)$ increases in time until it reaches a pole at the breakthrough when the expression in the brackets equals 0. If we introduce a constant hydraulic head, e.g. $h(\tau) = h_{\max}$ into Eq. (8), the breakthrough time equals:

$$T_B^{h_{\max}} = \frac{1}{2\gamma}\cdot\frac{n-1}{2n+1}a_0\cdot\frac{1}{F_{in}}\cdot\left(\frac{L}{\lambda_0}\right)^{\frac{n}{n-1}}. \tag{9}$$

Using Eq. (9) in Eq. (8) a more compact form of $a(t)$ is obtained:

$$a(t) = a_0\left[1-\frac{h_{\max}^{\frac{-n}{n-1}}\int_0^t h(\tau)^{\frac{n}{n-1}}d\tau}{T_B^{h_{\max}}}\right]^{\frac{1-n}{2n+1}}. \tag{10}$$

4.3. Mathematical assessment of the 'loop vs. canyon' question

The question of 'loop vs. canyon' has now been translated to a mathematical form. If $a(t)$ in Eq. (10) has a pole between $0 < t < t_2$, the result is a loop. The condition for a loop can thus be written as:

$$h_{\max}^{\frac{-n}{n-1}}\int_0^{t_2} h(\tau)^{\frac{n}{n-1}}d\tau \ge T_B^{h_{\max}}. \tag{11}$$

Integration over time-varying hydraulic head given in Eq. (1), gives:

$$\int_0^{t_2} h(\tau)^{\frac{n}{n-1}}d\tau = \left(\frac{h_{\max}}{t_1}\right)^{\frac{n}{n-1}}\int_0^{t_1} t^{\frac{n}{n-1}}dt + \left(\frac{h_{\max}}{t_2-t_1}\right)^{\frac{n}{n-1}}\int_{t_1}^{t_2}(t_2-t)^{\frac{n}{n-1}}dt = \frac{n-1}{2n-1}h_{\max}^{\frac{n}{n-1}}t_2. \tag{12}$$

The inequality (11) becomes:

$$\frac{n-1}{2n-1}t_2 \ge T_B^{h_{\max}}. \tag{13}$$

Dividing both sides by $T_B^{h_{\max}} = T\cdot h_{\max}^{n/(1-n)}$ (see Eqs. (9) and (5)) and applying relations for $h_{\max}$, t_1 and t_2 given in Eq. (1), the condition becomes:

$$\mathrm{LCR} = \frac{\frac{n-1}{2n-1}\cdot\Delta z^{(2n-1)(n-1)}\cdot\frac{1}{D}\cdot\left(1-\frac{D}{E}\right)^{n/(n-1)}}{T} \ge 1. \tag{14}$$

We have named the left-hand side of this inequality (14), 'LCR', a Loop-to-Canyon-Ratio. The higher the LCR, the higher the probability that a loop will be the result of the scenario. The condition demonstrates the importance of relatively high erosion rates of the valley with respect to the incision rate of the C1. This assures a necessary hydraulic head for the evolution of deeper pathway.

LCR has more than a quadratic dependence on Δz. When multiple competing pathways connect point A to the valley, this implies an advantage of deeper pathways, provided that all other parameters are similar for all pathways. However, we have neglected that the length of C2 also depends on Δz. Such an assumption is almost true when $L_H \gg \Delta z$. This result is also in accordance with some of the conclusions of Worthington (2004, 2005). If we assume that C2 is composed of a horizontal part with length L_H and two vertical parts with length Δz, its entire length can be approximated as $L = L_H + 2\Delta z$. The length is still combined with other parameters in the denominator T of the inequality (14). If we apply power law dependence of T on L, $T = \widetilde{T}\cdot L^{2n/n-1}$ as

it can be seen in Eqs. (9) and (5) assumes the fourth order kinetics ($n = 4$), the LCR can be rewritten in a more condensed version:

$$\mathrm{LCR} = \frac{\frac{3}{7}\Delta z^{7/3}\frac{1}{D}\left(1-\frac{D}{E}\right)^{4/3}}{(L_H + 2\Delta z)^{8/3}\widetilde{T}}. \quad (15)$$

In Eqs. (14) and (15), LCR is given as a function of parameters defining specific speleogenetic settings. Those characteristics for the scenario given in Fig. 3 (Δz, D, E and L_H) are explicitly written. All other parameters, such as the initial aperture of C2 and basic physical and chemical parameters and constants, are combined in $\widetilde{T} = T_B^{h_{\max}} \cdot h_{\max}^{4/3} \cdot L^{-8/3}$, with an explicit form given by:

$$\widetilde{T} = \frac{1}{3\gamma}\cdot\left(\frac{1}{a_0}\right)^3 \cdot \left(\frac{72\eta}{\rho\cdot g\cdot c_{eq}}\right)^{\frac{4}{3}} \cdot k_n^{1/3}. \quad (16)$$

Fig. 5a shows a dependence of LCR on Δz for D/E ranging from 0.9 to 0.1 in steps of 0.2. Full lines show results for $D = 0.005$ cm/y and dashed lines for ten times faster incision, $D = 0.05$ cm/y; $L_H = 1$ km, $a_0 = 0.02$ cm. The shaded area denotes LCR < 1, where a canyon is more likely to develop. A smaller inner graph shows the dependence of LCR on Δz for $L_H = 200$ m and $D/E = 0.5$. For $\Delta z \ll L_H$, LCR $\propto \Delta z^{7/3}$. LCR has a maximum at $\Delta z = 7/2 \cdot L_H$, and decreases for large Δz, asymptotically towards LCR $\propto \Delta z^{-1/3}$.

Fig. 5b shows a dependence of LCR on D/E for several Δz, ranging from 50 m to 10 m in steps of 10 m. Full lines show a scenario with $D = 0.005$ cm/y and $L_H = 500$ m and dashed lines for $D = 0.05$ cm/y and $L_H = 1$ km. Small D/E assures a driving hydraulic head for C2. Additionally, small D extends the time available for the formation of a loop.

Several approximations have been used to get the results given above, and therefore they have to be critically evaluated. The breakthrough time derived from such analytical solution is higher than the numerical results based on the finite difference calculations. However, the functional dependence on the basic parameters is the same (Dreybrodt and Gabrovsek, 2000, 2002).

5. Evolution after the breakthrough of the deep pathway

So far it has been assumed that immediately after the breakthrough, C2 takes all available recharge and C1 remains dry. One might question this assumption. In this section we assess the extent of growth of C1 after the breakthrough of C2. The penetration length along C2 becomes long enough to assume its uniform widening. Even more, the incision rate of C1 and the wall retreat in C2 will be taken as equal.

C1 receives water and evolves only if C2 is pressurised and cannot take all available recharge at point A (see Fig. 6). The head drop along C2 is given by the elevation difference between its input at point A and output in the valley. Taking the head difference and geometrical parameters known, the flow rate through C2 is given by the Darcy–Weisbach equation:

$$Q = \sqrt{\frac{2\cdot g\cdot D\cdot \Delta h}{f\cdot L\cdot A^2}} \approx \sqrt{\frac{4\cdot g\cdot a^3\cdot b}{f\cdot L}}. \quad (17)$$

f is a friction factor (about 0.05–0.1 for a fully turbulent flow in rough pipes), L the length of the pathway C2, V the flow velocity, D the hydraulic diameter, A the cross-section of the conduit and g the gravitational acceleration. The hydraulic diameter is defined as $D = 4A/P$, where P is a perimeter of a conduit. The expression on the right is given for a rectangular conduit with aperture width a and breadth b. For a wide conduit, $b \gg a$, an approximation $D \approx 2a$ is valid.

h_{TB} is then set as the elevation of A at the breakthrough of C2. If the base level is above the Level 2, the elevation of A from thereon increases as $h_{TB} + (E - D)\cdot t$, otherwise it decreases as $h_{TB} - D\cdot t$. In the latter case, the flow along C2 increases as:

$$Q(t) = \sqrt{\frac{4\cdot g\cdot b\cdot (a_{TB}+2Dt)^3\cdot (h_{TB}-Dt)}{f\cdot L}}. \quad (18)$$

Note that here the time t starts running at the breakthrough of C2. a_{TB} denotes the aperture width of C2 at its breakthrough, usually in the order of few millimetres.

C1 receives water until $Q_{in} > Q(t)$. Equating Q_{in} and $Q(t)$ and (reasonably) assuming that $z_{TB} \gg Dt$, gives the time available for incision of C1 after the breakthrough of C2:

$$t_{PB} = \frac{1}{2D}\left(\left(\frac{f\cdot L\cdot Q_{in}^2}{4\cdot g\cdot h_{TB}\cdot b}\right)^{\frac{1}{3}} - a_{TB}\right). \quad (19)$$

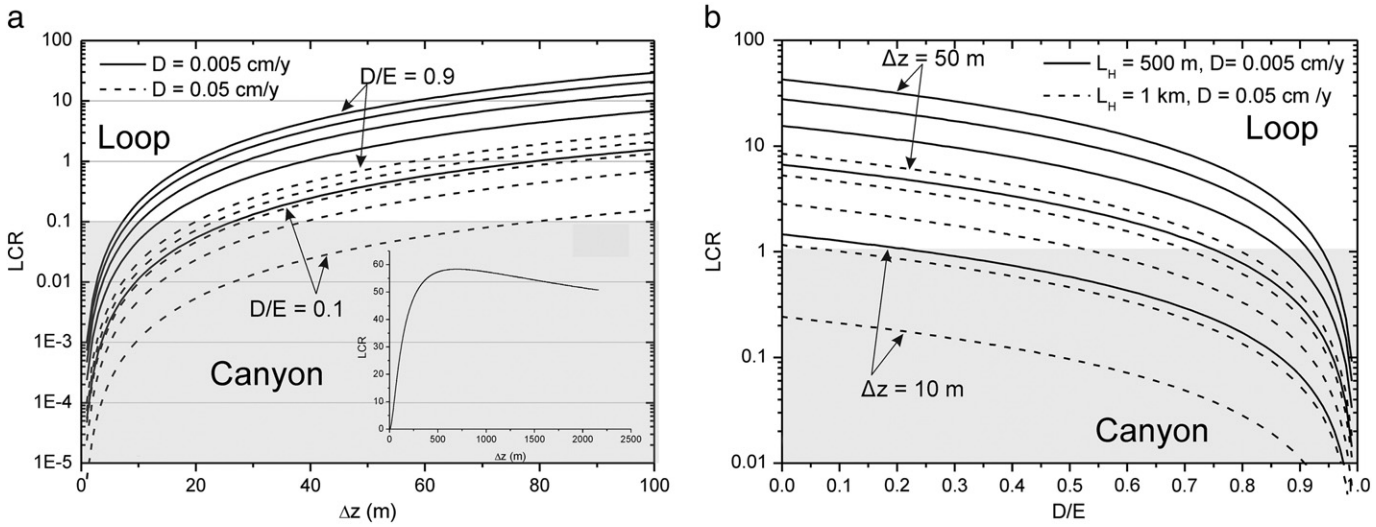

Fig. 5. a) Dependence of the LCR on Δz for values of $D/E = 0.9$ (top curve), 0.7, 0.5, 0.3 and 0.1 (bottom curve). Full curves represent cases for $D = 0.005$ cm/y and dashed curves for $D = 0.05$ cm/y. $L_H = 1$ km. Small inner graph shows the bigger range of LCR(Δz) for $L_H = 200$ m. b) Dependence of the LCR on D/E for values of $\Delta z = 50$ m (top curve), 40 m, 30 m, 20 m, and 10 m (bottom curve). Full curves represent cases for $L_H = 500$ m and $D = 0.005$ cm/y and dashed curves for $L_H = 1$ km and $D = 0.05$ cm/y.

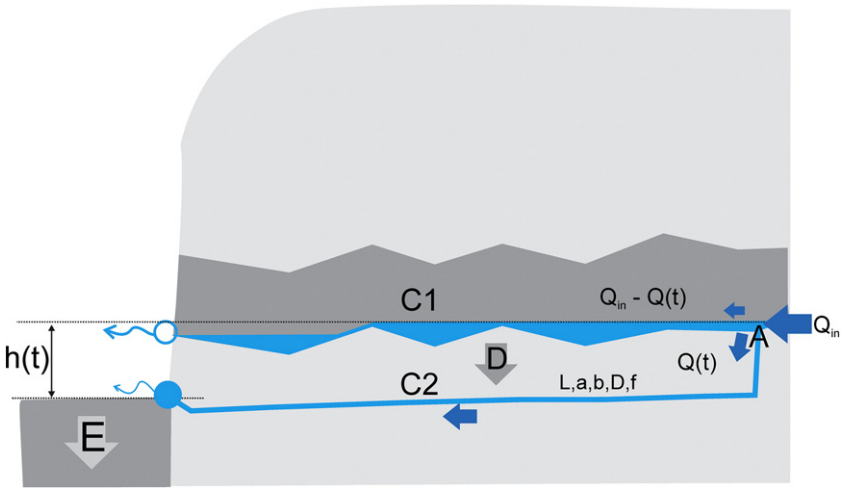

Fig. 6. Evolution of the system after the breakthrough of C2.

The total incision of C after the breakthrough of C2 is given by:

$$\Delta z_{PB} = t_{PB} \cdot D = \frac{1}{2}\left(\left(\frac{f \cdot L \cdot Q_{in}^2}{4 \cdot g \cdot h_{TB} \cdot b}\right)^{\frac{1}{3}} - a_{TB}\right). \quad (20)$$

Inserting a relevant range of parameter values (100 m < L < 1000 m, f = 0.1, 10 m < h_{TB} < 50 m, 1 m^3/s < Q_{in} < 10 m^3/s) results in t_{PB} ranging between several tens and several hundred years and Δz_{PB} from few centimetres up to a metre. This confirms that most of the canyon incision occurs before the breakthrough of the deep loop. Later deep pathways soon capture all the available flow. This confirms the assumption taken above, that the breakthrough time can be taken as a limiting time for the canyon incision.

6. The role of a transient recharge

A simple transient recharge scenario can be composed of a series of flood events, separated by periods of low water. The conduits are rarely straight, and they generally undulate or zig-zag along preferential pathways of their formation, forming crests and troughs. This observation, also denoted in Fig. 3, presents a key to the role of transient recharge. Fig. 7 shows two snapshots of the system during the flood event (Fig. 7a) and during the low water conditions (Fig. 7b). During the high water regime, C1 and C2 evolve as described in the previous section. During the low water, C1 ceases to incise as there is no permanent stream flowing along it. However, the water trapped in the troughs of C1, still feeds and widens deeper loops. This gives them an important advantage. If α denotes the time share of high water, when C1 is flooded, then the average incision rate of C1 is αD. To account for the effect, D in all of the above equations (e.g. Eq. (14)) can be replaced by αD. This increases the LCR by more than a factor of $1/\alpha$. In real systems, α can be in the order of 10^{-2}–10^{-1}, so transient recharge increases the chance for a loop for one to two orders of magnitude.

One might argue that the volume of water captured in troughs is too small to contribute to the evolution of loops. To the contrary, Fig. 4 demonstrates that only several tens of litres of water have passed through the fracture until the breakthrough.

Fig. 7 shows several sub-vertical pathways (soutirages) connecting Level 1 to Level 2 as initial parts of several competing loops. If differences of initial apertures are not concerned, the loop along S1, which has the highest hydraulic gradient, evolves most efficiently. After its breakthrough, the evolution along S2 is enhanced and later on followed by S3 and so on into the interior of a massif.

7. Discussion and conclusions

Although the presented results are based on several simplifying assumptions and generalisations, we believe that they have important implications for understanding speleogenesis in tectonically active environments with high uplift and erosion rates.

The initial hypotheses given in Section 2 were extended by stressing and quantifying the importance of base-level down-cutting. During valley incision, the drainage network in karst is continuously accommodating to

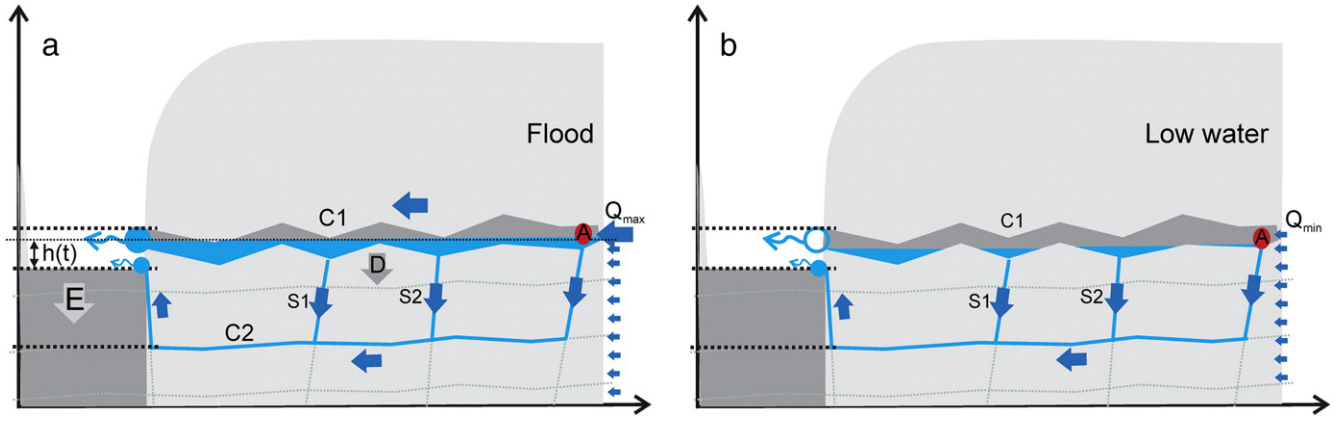

Fig. 7. a) The system during a flood event. b) The system during low water. Several sub-vertical pathways ('soutirages') drain water trapped in the troughs of C1 towards the spring.

the new boundary conditions with a water table cave as an end member. When erosion rates of a valley and incision rates of a karst channel are similar, the channel evolves into a canyon. If the valley deepens quickly compared to channel incision, hydraulic gradients build up and accelerate the evolution of looping pathways below the incising channel. There is a competition between formation of a canyon and evolution of looping pathways; the latter evolve more efficiently when valley erosion rate is fast compared to the incision rate of an overflow channel.

The nature of the recharge can play an important role in this competition. Flood water, stored in troughs of an upper main channel (i.e. C1), drains through sub-vertical pathways (soutirages) to a lower level during low flow. This extends the evolution of looping pathways and shifts the competition in favour of a looping cave.

In an ideal setting, where fracture distribution is relatively uniform with fast valley deepening, there is only a small chance for the development of water table conduits, as the evolution of deeper pathways is too fast. Natural karst areas are structurally and lithologically non-uniform, and the down-cutting of valleys is not continuous. Such irregularities still enable the concentration of conduit development along some levels. However, many examples from the Alps demonstrate that conduit formation shifts to deeper levels before a true end-member water table cave can form.

We have shown that formation of water table or looping caves is not principally dependent on fracture density but also on the recharge dynamics, valley incision rate and vertical distribution of permeable structures. In this sense, our findings offer new interpretation of field observation, which goes beyond that of the four state model of Ford and Ewers (1978).

One could consider C1 as a surface canyon. Therefore, all results and conclusions presented above are also valid for the competition between incision of a surface canyon and underground flow pathways.

We believe that this paper answers and opens some important questions related to the vertical organisation of karst networks. The discussion is based on a relatively simple analytical model, which enables some general conclusions. However, to extend it into the direction of possible resulting network geometries, one would have to apply more realistic 2D and 3D numerical models.

Acknowledgements

We would like to dedicate this article to professors Derek Ford and Wolfgang Dreybrodt. The four state model of Ford and Ewers was a great step forward in explaining speleogenesis for its time and was the initial inspiration for this work. The approach taken in this work is based on quantitative, process-based approach in speleogenesis, pioneered by W. Dreybrodt. The present article was partly inspired by fruitful discussions with Derek Ford, Arthur Palmer and Pierre-Yves Jeannin. Elaboration of the article was supported by the 'Proteus' governmental exchange programme 28079WE between France and Slovenia to Ph.A. and F.G. Work of F.G was supported by the Slovenian Research Agency, Project J2-4093. Warm thanks are extended to Arthur Palmer and David Culver for correcting the English of the original and revised manuscript, respectively.

References

Audra, Ph., 1994. Karsts alpins — Genèse de grands réseaux souterrains. Karstologia Mém. 5 (in French).
Audra, Ph., Palmer, A.N., 2013. The vertical dimension of karst — controls of vertical cave pattern. In: Shroder Jr., J., Frumkin, A. (Eds.), Treatise on Geomorphology, vol. 6. Academic Press, San Diego, CA, pp. 186–206.
Bögli, A., 1980. Karst Hydrology and Physical Speleology. Springer Verlag, Berlin.
Bretz, J.H., 1942. Vadose and phreatic features of limestone caverns. J. Geol. 50, 675–811.
Davis, W.M., 1930. Origin of limestone caverns. Geol. Soc. Am. Bull. 41, 475–628.
Dreybrodt, W., 1988. Processes in Karst Systems: Physics, Chemistry, and Geology. Springer-Verlag, Berlin.
Dreybrodt, W., 1996. Principles of early development of karst conduits under natural and man-made conditions revealed by mathematical analysis of numerical models. Water Resour. Res. 32 (9), 2923–2935.
Dreybrodt, W., Gabrovsek, F., 2000. Dynamics of the evolution of a single karst conduit. In: Klimchouk, A., Ford, D.C., Palmer, A., Dreybrodt, W. (Eds.), Speleogenesis: Evolution of Karst Aquifers. National Speleological Society, pp. 184–193.
Dreybrodt, W., Gabrovsek, F., 2002. Basic processes and mechanisms governing the evolution of karst. In: Gabrovsek, F. (Ed.), Evolution of Karst: From Prekarst to Cessation. ZRC Publishing, Ljubljana, pp. 115–154.
Dreybrodt, W., Gabrovsek, F., Romanov, D., 2005. Processes of Speleogenesis: A Modeling Approach. Zalozba ZRC, Ljubljana.
Eisenlohr, L., Meteva, K., Gabrovsek, F., Dreybrodt, W., 1999. The inhibiting action of intrinsic impurities in natural calcium carbonate minerals to their dissolution kinetics in aqueous H_2O–CO_2 solutions. Geochim. Cosmochim. Acta 63 (7–8), 989–1001.
Ford, D.C., 1971. Geologic structure and a new explanation of limestone cavern genesis. Trans. Cave Res. Group Great Brit. 13, 81–94.
Ford, D.C., 1999. Perspectives in karst hydrogeology and cavern genesis. Proceedings of the Symposium on Karst Modeling. Karst Waters Institute Special Publication, 5, pp. 17–29.
Ford, D.C., Ewers, R.O., 1978. The development of limestone cave systems in the dimensions of length and depth. Can. J. Earth Sci. 15, 1783–1798.
Ford, D.C., Williams, P., 1989. Karst Geomorphology and Hydrology. Chapman & Hall, London.
Gabrovsek, F., 2000. Evolution of Early Karst Aquifers: From Simple Principles to Complex Models. Zalozba ZRC, Ljubljana.
Gabrovsek, F., Dreybrodt, W., 2001. A model of the early evolution of karst aquifers in limestone in the dimensions of length and depth. J. Hydrol. 240 (3–4), 206–224.
Häuselmann, Ph., Jeannin, P.-Y., Monbaron, M., 2003. Role of epiphreatic flow and soutirages in conduit morphogenesis: the Bärenschacht example (BE, Switzerland). Z. Geomorphol. 47 (2), 171–190.
Jeannin, P.-Y., Bitterli, T., Häuselmann, Ph., 2000. Genesis of a large cave system: the case study of the North of Lake Thun system (Canton Bern, Switzerland). In: Klimchouk, A., Ford, D.C., Palmer, A.N., Dreybrodt, W. (Eds.), Speleogenesis: Evolution of Karst Aquifers. National Speleological Society, Huntsville AL, pp. 338–347.
Kaufmann, G., 2003. Modelling unsaturated flow in an evolving karst aquifer. J. Hydrol. 276 (1–4), 53–70.
Kaufmann, G., Dreybrodt, W., 2007. Calcite dissolution kinetics in the system $CaCO_3$–H_2O–$CaCO_3$ at high undersaturation. Geochim. Cosmochim. Acta 71 (6), 1398–1410.
Lauritzen, S.E., Ive, A., Wilkinson, B., 1983. Mean annual runoff & the scallop flow regime in a subarctic environment. Trans. Br.Cave Res. Assoc. 10 (2), 97–102.
Lismonde, B., Frachet, J.M., 1978. Grottes et scialets du Vercors (tome I & II). Comité départemental de spéléologie Isère, Grenoble.
Martel, E.A., 1921. Nouveau traité des eaux souterraines. Delagrave, Paris.
Palmer, A.N., 1991. Origin and morphology of limestone caves. Geol. Soc. Am. Bull. 103, 1–21.
Palmer, A.N., 2000. Digital modeling of individual solution conduits. In: Klimchouk, A., Ford, D.C., Palmer, A.N., Dreybrodt, W. (Eds.), Speleogenesis: Evolution of Karst Aquifers. National Speleological Society, Huntsville AL, pp. 194–200.
Swinnerton, A.C., 1932. Origin of limestone caverns. Geol. Soc. Am. Bull. 43, 662–693.
White, W.B., 2000. Development of speleogenetic ideas in the 20th century: the modern period, 1957 to the present. In: Klimchouk, A., Ford, D.C., Palmer, A.N., Dreybrodt, W. (Eds.), Speleogenesis: Evolution of Karst Aquifers. National Speleological Society, Huntsville AL, pp. 39–43.
Worthington, S.R.H., 1991. Karst Hydrogeology of the Canadian Rocky Mountains. (PhD thesis) McMaster University, Hamilton, Canada.
Worthington, S.R.H., 2004. Hydraulic and geological factors influencing conduit flow depth. Cave Karst Sci. 31, 123–134.
Worthington, S.R.H., 2005. Evolution of caves in response to base-level lowering. Cave Karst Sci. 32, 3–12.