THE KARST PARADIGM: CHANGES, TRENDS AND PERSPECTIVES
KRAŠKA PARADIGMA: SPREMEMBE, TRENDI IN PERSPEKTIVE

Alexander Klimchouk

Abstract

The paper examines representative definitions of karst (21), and discusses some concepts that influenced the modern understanding of the phenomenon. Several trends are discussed that took karst science beyond the limits of the traditional paradigm of karst. Dramatic progress in studies of speleogenesis plays the most significant role in changes taking place in the general understanding of karst. Also important is an adoption of the broad perspective to karst evolution which goes beyond the contemporary geomorphologic epoch and encompasses the entire life of a geological formation. Speleogenesis is viewed as a dynamic hydrogeological process of self-organization of the permeability structure in soluble rocks, a mechanism of the specific evolution of the groundwater flow system. The result is that these systems acquire a new, “karstic”, quality and more complex organization. Since almost all essential attributes of karst owe their origin to speleogenesis, the latter is considered as the primary mechanism of the formation of karst. Two fundamental types of speleogenesis, hypogene and epigene, differentiate mainly due to distinct hydrodynamic characteristics of the respective groundwater flow systems: (1) of layered aquifer systems and fracture-vein flow systems of varying depths and degrees of confinement, and (2) of hydrodynamically open, near-surface unconfined systems. Accordingly, two major genetic types of karst are distinguished: hypogene and epigene. They differ in many characteristics, notably in relationships with the surface, hydrogeological behaviour, groundwater quality, and the areas of practical importance and approaches to solving karst-related issues. Although views on essential attributes of karst have been clearly changing, this was not reflected in definitions of the notion which are in broad use in the earth-science literature. A refined approach is suggested to the notion of karst in which it is viewed as a groundwater (fluid) flow system of a specific kind, which has acquired its peculiar properties in the course of speleogenesis.

Keywords: karst, definition of karst, speleogenesis, karst geosystem, karst evolution.

Izvleček

Dokument obravnava reprezentativne definicije krasa (21) in obravnava nekatere koncepte, ki so vplivali na sodobno razumevanje tega pojava. Razpravlja o več trendih, ki so popelejali krasoslovje preko meja tradicionalne paradigme krasa. Dramatični napredek v speleogenetskih raziskavah je igral najpomembnejšo vlogo pri spreminjani, ki so se zgodile pri splošnem razumevanju krasa. Prav tako je za razvoj krasoslovja pomembno sprejetje široke perspektive, ki presega sodobno geomorfološko epoho in zajema celoten geološki razvoj. Speleogeneza je predstavljena kot dinamični hidrogeološki proces samoorganizacije stopnje prepustnosti posameznih struktur v topnih kamninah, to je kot mehanizem specifičnega razvoja sistema toka podzemne vode. Posledica tega je, da ti sistemi potrebujemo novo, ”kraško”, kakovostno in bolj kompleksno organizacijo. Ker skoraj vse bistvene značilnosti krasa izvirajo iz speleogeneze, je slednja šteta kot primarni mehanizem za nastanek krasa. Dve temeljni vrsti speleogeneze, hipogena in epigena, se razlikujeta predvsem zaradi različnih hidrodinamičnih značilnosti posameznih sistemov toka podzemne vode: (1) plavovodnosestni sistem in tokovni sistem po prelomih in žilah različnih globin in stopenj zapritosti, ter (2) hidrodinamično odprti in plitvi sistemi. Posledično razlikujemo dva velika genetska tipa krasa: hipogeni in epigeni. Med seboj se razlikujejo v mnogih značilnostih, predvsem v odnosih s površjem, po hidrogeološkem obnašanju, kakovosti podzemne vode, ter na področjih praktičnega pomena in pristopov k reševanju vprašanj, povezanih s krasom. Čeprav so se stališča o bistvenih lastnostih krasa spremenjala, se to ni odradilo v opredelitvi pojma, ki se na splošno uporablja v vedah o Zemlji. Pri predstavah o krasu predlagamo dodeljan pristop, v katerem je poudarek na sistem podzemne vode (tekočine) s specifičnim načinom pretakanja, ki je pridobil svoje značilne lastnosti prav tekom speleogeneze.

Ključne besede: kras, definicija krasa, speleogeneza, kraški geosistem, evolucija krasa.

1 Institute of Geological Sciences of the National Academy of Sciences of Ukraine, Ukrainian Institute of Speleology and Karstology, e-mail: klim@speleogenesis.info

Received/Prejeto: 07.09.2015
INTRODUCTION

The development of theoretical knowledge in a scientific discipline is uneven throughout its history. Acquisition of empirical data in karstology throughout the most of the 20th century outpaced their theoretical comprehension, and construction and integration of a conceptual superstructure. The development of the theoretical foundation of karstology occurred spontaneously, remaining largely within the scope of empirical generalizations. Certain conceptual models, sometimes successfully developed to the status of specific theories, remain poorly coordinated and harmonized within the overall theoretical body of karstology – its categorial structure and conceptual system. Since terminology is the reflection of conceptual constructions, dissonance and discrepancy in concepts promotes ambiguous and confusing usage of terms. In a broad sense, the problem of terminology cannot be resolved through conventions and glossaries; true improvements come through harmonization and integration of notions and concepts that stay behind terms.

The situation in karstology is further complicated by differences in the development paths of national and regional scientific schools in the study of karst, – North-American, Western-European, Eastern-European, that of the ex-USSR, etc., – which relied on their own methodological and scientific-philosophical traditions. Expressed particularities of karst in various regions also contributed to discrepancy between scientific schools and discordance in concepts and ideas.

These divisions, however, have been considerably smoothed during the last three decades in the process of integration of the global karst and cave science. Moreover, the diversity in notions and approaches brings some advantages, when it comes to synthesis and generalization. The particular role in the integration has been played by publication of outstanding textbooks (e.g. White 1988; Ford & Williams 1989, 2007; Palmer 2007), international monographs and collections of papers (e.g. Bosak et al. 1989b; Klimchouk et al. 2000; Andreo et al. 2010; Frumkin & Shroder 2013), encyclopedia (e.g. Gunn 2004; Culver & White 2005; White & Culver 2012) and ISI-indexed international thematic journals (Acta Carsologica, Journal of Cave and Karst Studies, International Journal of Speleology), as well as by the activity of international karst-related scientific bodies (IAH Karst Commission, IGU Karst Commission, UIS scientific commissions, IGCP projects, etc.), which held numerous broad-scope scientific events (e.g. International Speleological Congresses) and thematic symposia.

Nevertheless, significant uncertainty and discordance remains in the understanding of some basic notions of karst science, such as the notion of karst which denotes the principal object (entity) of this scientific discipline. Unsolved problems in the understanding of the essence of karst are getting more obvious and acute due to several ongoing circumstances:

- Rapid development of karst research in various aspects and increase in scientific and practical importance of karst knowledge. This leads to the wide recognition of karstology as a self-standing discipline of the Earth Sciences (the establishment of the Karst Division in the GSA in 2014 is one of the recent indications), but also highlights existing methodological pitfalls;
- Dramatic geographical “expansion” and geological “deepening” of data about karst, as well as methodological diversification of studies, have led to improved understanding of the variety of karst manifestations and characteristics, and of natural environments of karst development;
- Intense development of hypogene karst research has changed ideas about distribution of karst in the Earth’s crust and the range of lithologies in which karstification is possible;
- The ongoing process of globalization and integration of karst science, as well as rapid development of specific conceptual models and theories in karstology, reveals and collides differences in the views practiced by national, regional and discipline-specific schools.

Some trends in karst studies show obvious signs of the ongoing shift in the scientific paradigm of karstology. Although the views on the essence (essential attributes) of karst are clearly changing, this is not reflected appropriately in definitions of the notion which are in broad use in the earth-science literature. The purpose of this paper is to highlight important changes through discussing definitions of karst and some recent developments, and eventually to outline an approach to refinement of the notion of karst from the modern perspective.

“...from today's viewpoint it is unsatisfactory to regard the reference to the semantic origin of the term as correct conceptual definition of the karst in general” (Jakucs 1977, p.15)
VARIANTS OF DEFINITIONS OF KARST

THE ORIGIN OF THE TERM “KARST”

The notion of “karst” owes its origin to the geographic area in Slovenia (Kras), which Germanized name was adopted as a scientific term. The wider area of SW Slovenia, between Ljubljana, Trieste and Rijeka, characterized by irregular barren rocky ground with karrens, dolines, deep gullies, sinking rivers, large springs, poljes, and caves is considered as the “Classical Karst.” In early studies of the Classical Karst and similar areas around Europe and elsewhere, the emphasis was placed on most obvious, readily observable and impressive geomorphological and hydrological aspects. Caves, although noted and recognized as a characteristic attribute of karst quite early, were long regarded as a curious phenomenon rather than objects of prime scientific importance and systematic studies. Their paramount role (more broadly, - the role of the conduit permeability) in the development of karst has been explicitly appreciated only during last four decades.

PHENOMENA IN CARBONATE ROCKS

One of the corollaries of such origin of the term is an attribution of karst exclusively or mainly to carbonate rocks, a trend that started from the Martel’s definition of karst as phenomena in limestones (Martel, 1984) and deeply rooted in the literature. This is illustrated by definitions given below.

[1] Karst is a landscape formed upon and within carbonate rock sequences by the dissolitional effect of carbonic acid (Lowe 1992).
[2] Karst is primarily a landscape, with specific landforms and solution features, which are mainly developed in carbonate rocks (Cost Action 65 1995).
[4] Karst is defined as a limestone landscape with underground drainage (Lahr 2003).
[5] Karst features mainly occur in carbonate rocks, limestone and dolomite, in which formations it is considered as true karst (Bakalowicz 2005).
[6] Karst: Landforms that have been modified by dissolution of soluble rocks (limestone and dolostone) (Poucher & Copeland 2006).

Some researchers further argued that only chemical dissolution mechanisms comprising three components in phase equilibrium, such as carbonic acid dissolution of carbonates, can be regarded as producing a true karst (karst sensu strict). Cigna (1978, 1985) suggested using the term parakarst for karst developing in gypsum (two components in phase equilibrium). Lowe (1992), having stated a similar viewpoint, referred to terms such as evaporite karst (including halite, gypsum karst and anhydrite karsts) as hybrid terminology.

Karst in evaporite rocks, especially in gypsum, widely occurs in many regions, particularly in the North America, Eastern Europe, Eastern Siberia and the Middle East. Despite of differences in dissolution mechanism, it results principally in the same set of phenomena that are attributable to carbonate karst, although distribution and appearance of evaporite karst has some specifics. Due to high solubility and dissolution rates, gypsum and salt rarely survive when exposed at the surface, but these rocks (especially gypsum) widely occur and get readily karstified in interstratal deep-seated settings, although such karst is often scarcely or not at all manifested at the surface. Much research had been done on evaporite karst in the ex-USSR, USA, Italy and some other countries, but its recognition in the mainstream karst science was somewhat slow until the mid 1980s. Since that, a series of thematic collections and monographs appeared (Authors varia 1986; Klimchouk et al. 1996; Johnson & Neal 1997; Calaforra Chordi 1998; Younger 2005; Gutiérrez et al. 2008), as well as hundreds of papers in international journals. This firmly established karst in evaporitic rocks as a part of the “true karst”, although acknowledging its expressed specifics due to particular features of geological occurrence and dissolution mechanisms. Karst in evaporites has been given due attention in major textbooks (Ford & Williams 1989, 2007) and international encyclopedia (Gunn 2004; Culver & White 2004, White & Culver 2012).

The extension of the arena of karst to a variety of readily soluble rocks other than carbonates already represents a considerable drift from the traditional “Classical Karst”-based understanding of karst. Even more far-reaching deviation is an extension of the karst concept to relevant phenomena in quartzites (Wray 1997), strongly supported by recent developments in hypogene karst studies (discussed below).

1 The prefix par implies something that is similar to the parent, but is not a true phenomenon.
KARST AS A SPECIFIC TOPOGRAPHY, LANDSCAPE, OR A SET OF LANDFORMS (PHYSIOGNOMIC APPROACH)

Another, and the most essential consequence of the geographic origin of the term and the early emphases in karst studies was the attribution of karst to a type of landscape. Statements that karst is a landscape (a specific type of landscape, or a set of landforms) form the core parts in most of the definitions given above [1, 2, 4, 6, 7] and in the following examples:

[8] *Karst is a special type of landscape that is formed by the dissolution of soluble rocks, including limestone and dolomite* (Karst Water Institute 2015).

[9] *Karst is a landscape formed from the dissolution of soluble rocks including limestone, dolomite and gypsum. It is characterized by sinkholes, caves, and underground drainage systems* (University of Texas at Austin 2015).

[10] *Karst - a terrain, generally underlain by limestone or dolomite, in which the topography is chiefly formed by the dissolving of rock, and which may be characterized by sinkholes, caves, and underground drainage systems* (Monroe 1972).

Some definitions [3] denote karst as a terrain with distinctive hydrology and landforms. These include the definition that is most widely accepted and cited in the modern literature:

[11] *Karst is terrain with distinctive hydrology and landforms arising from the combination of high rock solubility and well developed solutional channel (secondary) porosity underground* (Ford & Williams 1989; Ford 2004).

The definitions in this group present karst as a geomorphological (physiognomic) category, or a territorial entity, a certain part of the Earth’s surface characterized by specific features. Many definitions refer to distinctive features of landscape/terrain without specifying them or their specific properties, implying or stating that the distinctiveness (specifics) arises from the origin of landforms by dissolution. Some forms commonly seen as karstic are in fact created by heterogeneous processes with limited, hard-to-assess, or spatially remote roles of dissolution, but they are considered to be karstic if “... solution plays an essential precursor or ‘trigger role’” (Ford 1980, p. 345). Other definitions include lists of forms deemed to be characteristic for karst, though these lists are inevitably incomplete, arbitrary and biased. Present knowledge of the diversity of natural conditions/environments of karst development and respective diversity of karst manifestations further supports the view of Huntoon (1995) that dwelling upon ambiguous morphological character of karst is not a promising approach to unambiguously define it.

The definitions cited above are representative for the traditional karstological paradigm, clearly dominated by the geomorphological perspective in the understanding of karst. Within this paradigm, karstification is inalienably related to the surface and the meteoric recharge that comes from it, either diffused authogenic or concentrated alloigenic. The erosion base level ultimately controls the development of karst. Karstification commences either immediately after deposition and early exposure, or when combined action of uplift and denudation brings a buried soluble formation back to a shallow subsurface and re-exposure. This is a conceptual model of epigene karst, which dominated the whole karst science until recently. The terms and concepts such as covered karst, buried karst, palaeokarst, exhumed karst, etc. reflect the central role of the surface exposure in the traditional paradigm, which is further illustrated by a common but misleading belief among geologists that the presence of any karstified interval in a stratified sequence ultimately indicates an unconformity and a period of subaerial exposure.

However, in practice of karst studies, the departure from the above outlined traditional understanding of karst is old and massive. They originated early from the Cvijic’s merokarst (“imperfect” karst in impure limestones, with marly interbeds), and continued through the Maksimovich’s (1963) “Russian” type of karst (“zakryty” karst - “closed”, or “confined” karst that develops beneath the pre-karst insoluble cover) and Quinlan’s (1978) interstratal karst with largely the same meaning, to deep-seated karst in the evolutionary meaning of Klimchouk (1996) and Klimchouk & Ford (2000). See also the discussion of these and other terms in Bosak et al. (1989a) and Palmer & Palmer (1989). Furthermore, it is obvious that ideas of hydrothermal karst, sulfuric acid karst, artesian karst and hypogenic karst did not fit into the traditional karst paradigm. Thousands of recent publications deal with deep-seated karstification that was not formed, and is not manifested at the surface. Of importance here is that the notion of karst has clearly decoupled from the attribute of a distinctive suite of surface landforms, which in fact is not the required one.

Confined in the geological but not necessarily in the hydro-geological sense.
KARST AS A GEOLOGICAL ENVIRONMENT
Huntoon (1995) viewed karst as a geologic environment with specific properties:

[12] Karst is a geologic environment containing soluble rocks with a permeability structure dominated by interconnected conduits dissolved from the host rock which are organized to facilitate the circulation of fluid in the downgradient direction wherein the permeability structure evolved as a consequence of dissolution by the fluid (Huntoon 1995).

This definition indicates further departure from the terrain/landscape-based understanding of karst. It points to several important properties, which have been missed in most other formulations, such as conduit-dominated permeability structure, its organisation, and circulation of fluid. These and other properties are further discussed below.

KARST AS A TOTALITY OF PHENOMENA (FORMS)
Some definitions denote karst as a totality of phenomena or forms in soluble rocks. These include the already mentioned Martel’s understanding of karst as phenomena in limestones and are further illustrated by the following examples:


These definitions are obviously too inclusive; with the absence of specific designators they allow inclusion of phenomena and forms of various natures that may be present in soluble rocks.

KARST AS A PROCESS
A process-based approach to defining karst long dominated in the ex-USSR, with the following definition being most widely accepted:

[15] Karst is a process of chemical and partly mechanical action of underground and surface non-stream waters upon soluble permeable rocks (Maksimovich 1963).

[16] Karst is a process of destruction and obliteration of permeable soluble rocks mainly through their leaching by moving water (Sokolov 1962).

[17] Karst is a heterogeneous process of interaction of rocks and underground waters, that is dissolution of the former and removal of dissolved rocks by the latter (Zverev 1999).

[18] Karst – a "throughgoing" process of metasomatic alteration of rocks, which consists of the creation and subsequent filling of cavities and proceeds along a general scheme: dissolution-transport-precipitation of the matter (Ezhov et al. 1992).


Definitions [15] and [17] present karst as a process of water-rock interaction with the leading role of dissolution. Maksimovich reduced the process to chemical and mechanical action of water upon the rock, but the Sokolov’s definition [16] is an example of a broader view of karst as a geologic process. Ezhov et al. (1992) were first to suggest an original and promising view of karst as a metasomatic (alteration) process [18]. Also very original was the view of karst by Devdariani [19], who emphasized properties such as the self-development and flow concentration, which at the time were not acknowledged by other researchers as essential ones.

KARST AS A UNITY OF THE PROCESS AND FORMS
Some workers understood karst as a unity of the process and the resulting forms:


Such amalgamation of different categories seems to be methodologically questionable: the definable real-world entity must have a distinct categorial essence.

KARST AS A SYSTEM
The systems approach to karst (consideration of the object as system of interconnected components / properties / relations that constitute the complete whole) was emphasized and realized in some studies (e.g. Ford & Williams 1989), although it was not implemented in the definition of the notion of karst. Using in some definitions of wording such as "Karst is a system of…" remained declarative because the lack of clear identification of the object-system and indications of system-forming and emergent properties, for example:

[21] Karst is a system of processes and phenomena arising and developing underground and at the surface as the result of interaction of natural waters with rocks soluble in a given environment (Andreychouk 1991).
This brief overview of definitions demonstrates substantial discordance in the understanding even of a categorial status of karst. It also shows substantial deviations from the original landscape/terrain based idea of karst (physiognomic, geomorphological) towards process-based, geosystemic, and geological (hydrogeological) notions. For further discussion, it is necessary to look at some important developments in karstology which influence significantly the current general understanding of karst.

THE SUPREMACY OF GROUNDWATER FLOW IN SPELEOGENESIS AND KARST DEVELOPMENT

The notion of groundwater flow is used in this account in a sense of the “water exchange” term, commonly used in the Russian-language literature, which denotes a process characterizing recharge of the groundwater flow system, movement within it, and outflow to adjacent systems (discharge) (Shestopalov 1989). The groundwater flow system is a system of underground waters that is characterized by a common drive for movement and common conditions for circulation (Kartzev 1972).

Speleogenesis (karstic) is essentially a coupled mass-transfer / mass-transport process, which depends on both, the aggressiveness of groundwater and its flow. Flow (movement) is an inalienable attribute of groundwater. Aggressiveness results from disequilibrium in the water-rock system, and is an attribute of the moving groundwater (but not the opposite). Hence, the groundwater flow is the main controller of the equilibrium/disequilibrium state of the water-rock system, the main reason for the inaccessibility of equilibrium (achieved only locally and temporarily), which the system is seeking. It is a systematic transport and distribution mechanism that produces and maintains the disequilibrium conditions (Toth 1995; Shvartzev 2005).

To cause the speleogenetic development, dissolution effects of disequilibria have to accumulate over sufficiently long periods of time and/or to concentrate within relatively small rock volumes or areas. In a given rock media, the character of the water-rock interaction and the distribution of its effects are determined by the nature, intensity and pattern of the groundwater circulation, i.e. by hydrodynamic characteristics of the groundwater flow systems.

The above paragraphs provide a fundamental reasoning why the essence and principal categories of speleogenesis, and of karst in general, should be examined and established primarily based on the hydrogeologic perspective.

SPELEOGENESIS AS A PRIMARY PROCESS IN THE FORMATION OF KARST

Although caves were recognized as a characteristic attribute of karst in early karst studies, and the origin of caves has been a subject of lively debates since the beginning of the 20th century (Lowe 2000; White 2000), karst studies in general were long decoupled from the knowledge of caves. The role of speleogenesis in general understanding of karst, both in karstology and in the mainstream geosciences, was not properly acknowledged until recently. One of the problems was the domination of the anthropocentric notion of caves as human-enterable cavities – many areas displaying karstic physiognomy lack such caves. Another reason was that cave explorations became massive and pervasive, allowing to grasp a true scale of the phenomena, only after the mid of the 20th century, and comprehension by karst scientists of the enormous body of new data acquired by explorers is still in progress. And the third reason is that the theory of speleogenesis achieved its maturity and the ability to seriously influence other branches of geosciences only by the end of the last century. It is argued in this paper that it is the dramatic progress in studies of speleogenesis that plays the most significant role in the changes taking place in the general understanding of karst.

Although observations in human-enterable caves greatly contributed to the development of speleogenetic studies, the concept of speleogenesis is meaningless with respect to caves in the anthropocentric connotation of the term (Ford & Williams 1989). Such caves constitute fragments of natural void-conduit systems, which
geometry and spatial position are artificially, accidentally and vaguely defined. In studying speleogenesis, we always mean not the origin of such very fragments but the origin of void-conduit systems in their functional and structural integrity. Moreover, the major problems in the origin of caves lie in their inception and early development, which leaves no room for any anthropocentric scales in defining the study object. The term “cave” (karstic) is used in most speleogenetic studies and in this account to denote conduits and voids that are substantially enlarged, as compared to unmodified fractures and pores in the host rock, due to mainly dissolutional removal of the matter by fluid flow. It is commonly believed that apertures of conduits constituting cave systems start from a few mm but they can reach many meters in the course of speleogenesis.

A chain of developments that led to the modern understanding of speleogenesis began with the empirical generalization by Ford (1971) on the relationship between the evolution of the cave and the water table or piezometric surface (the “Four-State model”, which provided the resolution of what was called “a central problem of cave origin” during preceding decades), innovative physical modeling by R.O.Ewers of the evolution of epigene cave patterns (Ford & Ewers 1978; Ewers 1982), and recognition by White (1977) of implications for the early conduit development of an abrupt drop in the rates of calcite dissolution that occurs when the solution approaches chemical equilibrium (Berner & Morse 1974; Plummer & Wigley 1976). This kinetic change enables enlargement of initial flow pathways, although at slow rates, over long distances. When the initial pathway (a proto-conduit) is enlarged to the point where water is able to penetrate its entire length still retaining the substantial degree of undersaturation, the conduit growth rates accelerate dramatically, reaching 0.01-0.1 cm/yr, due to the positive feedback between increasing flow and dissolution (Dreybrodt 1990, 1996; Palmer 1991). This moment is termed “breakthrough”, and it signifies the birth of a cave (karst conduit). Importantly, it commonly coincides roughly with the thresholds marking the transition to turbulent flow regime and the onset of sediment transport (Ewers 1973; White 1977). The initial breakthrough, rapid enlargement of one or a few conduits and a drop in head in them are followed by drastic changes in the gradient field, re-organization of flow in the aquifer and cascading breakthroughs in tributary proto-conduits, - the process leading to the creation and increasing integration of conduit networks.

Much of the progress has been achieved through development and employment of numerical models that combine hydrodynamics with dissolution kinetics. This route has been paved by Dreybrodt (1990) and Palmer (1991) and advanced by many other works published during last 25 years, reviewed and summarized in Dreybrodt et al. (2005). These works confirmed some basic principles established by earlier empirical and physical modeling studies (Ford 1971; Ewers 1973, 1982; Ford & Ewers 1978; Ford & Williams 1989; Palmer 1991) and dramatically deepened our knowledge on how individual conduits and the patterns of conduits evolve depending on various boundary conditions and variables.

These studies revealed important general regularities in the evolution of conduit networks and highlighted the role of speleogenesis in the formation of karst aquifers, and more generally – of karst. Speleogenesis is driven by the positive feedback between discharge and dissolutional removal along initial flow pathways, and commonly includes three phases: (1) early speleogenesis (proto-speleogenesis) – slow widening of initial flow pathways, the development of proto-conduits; (2) speleogenetic initiation – the cascading process of breakthrough of the proto-conduits, which is characterized by their strong hydrodynamic competition for flow and increase in the growth rates, destabilization and re-organization of the flow field, transformation of boundary conditions, the emergence of integrated conduit systems and the formation of the contrasting level of conduit permeability; (3) speleogenetic development – stabilization of the system in a state of dynamic equilibrium by increasing the energy exchange with the environment, and further growth of conduits.

It should be noted here that the mechanism of formation of karst conduits that includes the flow-dissolution feedback, the achievement by proto-conduits of the regime of rapid dissolution kinetics (now called ‘breakthrough’) and further hydrodynamic competition and concentration of flow, was conceptually described by Lukin (1966) well before the discovery of the kinetic threshold and later modeling works. These ideas, however, did not receive due attention and further elaboration at the time, but they were perfectly confirmed by later studies.

Numerical modeling studies show that, regardless of the initial permeability structure of soluble rocks, speleogenetic evolution leads to the formation of a new, and the most contrasting level of porosity and permeability – an integrated system of conduits with apertures above the millimetric scale. This underscores the significance of a concept of multi-level porosity/permeability, originally introduced by Barenblatt et al. (1960) and first applied to karstified aquifers by Borevsky et al. (1973, 1976), now widely employed in characterization of karst aquifers. Importantly, patterns of void-conduit
systems created by speleogenesis are clearly organized to facilitate the most efficient groundwater flow in the downgradient direction, and physical and numerical models have revealed details of the self-organization process.

As the result of the speleogenetic evolution, the flow systems acquire important new properties that make them distinct, including (1) high heterogeneity and anisotropy of porosity and permeability, and (2) concentration of flow, both being the direct consequences and indications of self-organization (Huntoon 1995; Worthington & Ford 2009). Porosity comprised by integrated void-conduit systems is commonly low (within 0.05–3%), comprising only a small portion of total porosity of the rock media, but it provides for high hydraulic conductivity of aquifers (up to 1 m/s and higher) and transmit almost all (up to 99.9%) flow (Worthington et al. 2000). Additionally, the high degree of flow concentration is illustrated by a high proportion of large springs in regions underlain by soluble rocks, as compared to non-karstic regions, and the high efficiency of karstic flow systems is illustrated by very high flow velocities, commonly within $10^{-3}$–$10^{4}$ m/day (Worthington & Ford 2009), i.e. 5–7 orders of magnitude higher than typical velocities of groundwater movement in non-karstic flow systems in the zone of intense circulation.

Several theorizations, based on this brief review, can be made about the essence and the roles of speleogenesis.

Worthington & Ford (2009) have emphasized and reinforced the idea, previously expressed by Devdariani (1962, [19]) and Huntoon (1995, [12]) that speleogenesis is a specific process of self-organization of permeability and groundwater flow in soluble rocks. The speleogenetic initiation phase, i.e. the cascading process of the breakthroughs to the conditions of rapid growth (Ford 1980; Ford & Williams 1989), includes a radical re-organization of permeability structure of the flow system that eventually transforms boundary conditions and the functioning of the system.

Speleogenesis is a function of groundwater flow. Flow through soluble rocks inevitably leads to the formation of organized void-conduit systems (e.g. to speleogenesis), but speleogenesis, in turn, radically changes the structure and dynamics of flow systems. As noted by Ford & Williams (2007, p. 116), “… the karst circulation system undergoes more feed-back giving rise to continuous self-adjustment than occurs in any other type of groundwater system.” Therefore, speleogenesis can be viewed as a dynamic hydrogeological process of transformation of porosity and permeability structure of soluble rocks, as a mechanism of the specific evolution of groundwater flow systems, which results in that these systems acquire a new, "karstic", quality and the more complex and contrasting organization. This understanding emerged in a number of works through the 20th Century, but it was explicitly shaped by the 1990s (Ford & Williams 1989). It has been codified in the title and in the contributions of the major international monograph on the subject (Speleogenesis: Evolution of Karst Aquifers; Klimchouk et al. 2000) and strongly reinforced during the subsequent decade (Worthington & Ford 2009).

Since most of other (than hydrogeological) specific properties attributed to karst, including geomorphological ones, owe their origin to the development of organized dissolution porosity/permeability structures in soluble rocks (Palmer 1991), speleogenesis should be considered as the primary mechanism of the formation of karst. The onset of the speleogenetic initiation phase signifies the birth of karst. Therefore, in contrast to the view, tacitly implied within the traditional karst paradigm, that speleogenesis is the result of karst development, one can assert that the opposite is true, – karst is a function of speleogenesis.

SELF-ORGANIZATION OF PERMEABILITY AND FLOW SYSTEMS

Basic principles of self-organization in natural systems have been derived from non-equilibrium thermodynamics (Prigogine & Nicolis 1977; Prigogine 1980; Prigogine & Stengers 1984) and further developed by synergetics (Haken 1984, 2004), an interdisciplinary science explaining the formation and self-organization of patterns and structures in open systems far from thermodynamic equilibrium (dissipative systems). Besides physics and biology, these ideas and concepts received intense application in geology and geomorphology (Huggett 1988, 2007; Pozdnyakov & Chervanev 1990; Letnikov 1992; Phillips 1992; Gregory & Goudie 2011), geochemistry (Ortoleva 1994) and hydrogeology (Yakovlev & Borevsky 1994; Shvartzev 2005, 2008).

Self-organization is the spontaneous often seemingly purposeful formation of spatial, temporal, spatiotem-
poral structures or functions in systems composed of few or many components (Haken 2004). In dissipative systems, nonequilibrium is the source of order, with spontaneous fluctuations, that are amplified by positive feedback, growing into macroscopic patterns (Huggett 1988; Phillips 1992). The fluctuations trigger an instability that the system accommodates by reorganizing itself. Self-organization means an enormous reduction of degrees of freedom (entropy) of the system which macroscopically reveals an increase of “order” (pattern-formation). This far-reaching macroscopic order is independent of the details of the microscopic interactions of the subsystems (Haken 1984).

Groundwater systems are typical dissipative systems (Shvartzev 2008), and speleogenesis is an excellent example of self-organization of groundwater flow systems (Worthington & Ford 2009; Klimchouk & Andreychouk 2010). During the early speleogenesis phase, fluctuations in initial structural and chemical conditions of the water-rock interaction and the positive flow-growth feedback result in non-uniform development of proto-conduits. The emergence of first conduits (i.e. those in which breakthrough has occurred) destabilizes the system. The speleogenetic initiation phase is manifested in a series of cascading breakthroughs of proto-conduits to the outflow boundary and adjacent successful conduits, which causes further instability and continued transformation of fields of hydrodynamic and chemical parameters. In terms of synergetics, this phase is a giant fluctuation (bifurcation, or threshold) in the evolution of the open system, during which a new pattern emerges. Self-organization of the permeability structure (formation of an integrated conduit pattern) leads to transformation of boundary conditions and subsequent stabilization of the groundwater flow system at a new higher level of energy exchange with an external environment. The following speleogenetic development phase is a “stationary” stage, characterized by dynamic equilibrium.

As a result of speleogenetic self-organization, the groundwater flow system acquires a new, more complex structure and changes the functioning, i.e. it receives a new quality and can be attributed to a higher level of geosystemic organization.

TYPES OF KARST BEYOND THE TRADITIONAL PARADIGM, AND APPROACHES TO KARST TYPOLOGY

HYPOGENE KARST

Since hypogene karst during its formation is almost exclusively represented by voids and conduits, which origin is by definition unrelated to the surface agencies, the terms “hypogene karstification” and “hypogene speleogenesis” are virtually interchangeable. Ideas that karst can develop at depth without direct genetic relationship to the surface (i.e. without exposure of the host rocks and recharge from the immediately overlying surface) have a long history, but remained on the periphery of karstological thinking, not influencing the traditional paradigm of karst until the last 25 years.

Early scientific comments that solution cavities can form at depth due to the action of rising hydrothermal waters were made in the mid of 19th century by geologists who studied ore deposits in Europe. They went unnoticed by scholars of the first half of 20th century who shaped the body of the emerging science of karst. Since the mid of 20th century, ideas of hydrothermal karst, sulphuric acid karst and ore karst received further development mainly in Czech Republic, Hungary and the ex-USSR. In these countries, the concepts of deep-seated karstification driven by hydrothermal and sulphuric acid dissolution were easily fit into the process-based general notion of karst which was common there (see definitions [15] and [16]). Notable publications of this period include, among others, Jakucz (1948, 1977), Kunsky (1957), Sokolov (1962), Maksimovich (1969), Dublyansky (1980), Müller (1974), Sass-Gustkiewicz & Dzulynski (1982). In the United States, several important publications appeared, focused on speleogenesis by thermal waters (Egemeier 1973; Bakalowicz et al. 1987), sulphuric acid (Morehouse 1968; Hill 1987) and artesian waters (Brod 1964; Howard 1964). These speleogenetic works were clearly inconsistent with the landscape-based (epigenic) karst notion which dominated in the Western literature (see definitions [1–4 and 6–11]). Although these alternative mechanisms for cave development have been given due attention in the major contemporary text on karst (Ford & Williams 1989), this acknowledgement did not gain reflection in the approach to the general notion of karst ([11]).

The beginning of 1990s has been marked by several publications that signified the turning point in studies of hypogene speleogenesis. The book by Y.Dublyansky (1990) was the first comprehensive account on hydrothermal karst, including theoretical aspects. In his classical paper on the origin of limestone caves, Palmer (1991) has provided an excellent summary on hypogene
speleogenesis and brought the term “hypogone caves” into a broad international usage. Klimchouk (1990, 1992, 1994, 1997a) revitalized the concept of artesian speleogenesis by employing concepts of cross-formational communication in leaky confined aquifer systems. He demonstrated that giant gypsum maze caves in the Western Ukraine were formed by upward flow across the gypsum bed, sandwiched between two aquifers, in zones of topographic/piezometric lows. The small book by Ezhov et al. (1992) offered a thought-provoking and far-reaching discussion of “non-traditional” types of karst (hydrothermal karst, sulfuric acid karst, ore karst, silicate karst, endokarst, etc.) in the context of thermobaric conditions in the Earth’s crust. Being published in Russian by an obscure publisher, this important work was not properly appreciated even in Russia and remained unnoticed internationally.

Palmer (1995) has overviewed geochemical models for the origin of macroscopic solution porosity in carbonate rocks, and demonstrated a multiplicity of dissolution mechanisms operating in deep-seated mesogenetic environments. Klimchouk (2000) provided a lengthy review of speleogenesis in deep-seated and confined settings and relevant karst concepts, introduced the concept of transverse speleogenesis, highlighted the distinctiveness of deep-seated speleogenesis with respect to speleogenesis in unconfined settings and called for a revision and expansion of the traditional paradigm of karst in order to embrace the deep-seated phenomena. The multi-author international book on speleogenesis (Klimchouk et al. 2000) has codified the division of basic genetic settings for caves into (1) coastal and oceanic (eogenetic), (2) confined deep-seated (hypogenic), and (3) unconfined (hypergenic/epigenic).

By the end of the 20th century, the notion of the hypogene origin remained largely limited to caves formed by hydrothermal and sulfuric acid dissolution (Ford & Williams 1989; Palmer 1991; Hill 2000), and the term and concept of hypogene speleogenesis were linked to the origin (relative to the surface) of the aggressiveness of water (Palmer 1991). Klimchouk (2000) emphasized an importance for deep-seated speleogenesis of upwelling cross-communication between aquifers in leaky confined systems, and Ford (2006) suggested a definition of hypogene speleogenesis based on recharge from below. This approach, which can actually be traced from the recognition by Ford (1987) of a class of basal injection caves, has been further elaborated by Klimchouk (2007), who suggested that in hypogene speleogenesis the specific hydrogeological settings, including leaky confinement and upwelling flow pattern, transcend the particularities of the physico-chemical mechanisms which create the aggressiveness of water toward rocks. Therefore, hydrogeological criteria are decisive in distinguishing hypogene speleogenesis; this also follows from the general postulate of the supremacy of groundwater flow in speleogenesis (Section The supremacy of groundwater flow in speleogenesis and karst development). The author defines hypogene speleogenesis as the formation of solution-enlarged permeability structures (void-conduit systems) by fluids that recharge the cavernous zone from hydrostratigraphically lower units, being originated from distant, estranged (by low-permeability beds or strata), or internal sources, independent of direct recharge from the overlying or immediately adjacent surface (modified from Klimchouk 2007).

The hydrogeological approach highlights the common hydrogeological genetic background and explains the multifaceted similarity of caves formed by upwelling flow, previously seen as unrelated because of their attribution to different chemical processes involved. Importantly, it provides a theoretically and methodologically sound basis not only for defining and identifying hypogene speleogenesis, but also for its spatial and temporal prognosis in the context of regional hydrogeology and geodynamics (Klimchouk 2013b, 2013c, 2014).

In 1990s, independently from karst and cave science, sedimentologists and petroleum geologists studying carbonate reservoirs began to realize limitations of the model of subaerial meteoric diagenesis, heavily used to explain the formation of deep-seated dissolutional porosity in carbonates. This model implied that such porosity is related to past exposures and dissolution in paleo-vadose and paleo-phreatic freshwater zones (i.e., is paleokarst; Esteban & Wilson 1993). Some workers proposed that deep-burial dissolution in the mesogenetic environment can contribute significantly to secondary porosity and permeability evolution in many carbonate reservoirs (e.g., Mazzullo & Harris 1991, 1992; Al-Shaieb & Lynch 1993; Machel 1999). It was shown that mesodiagenetic dissolution in carbonate reservoirs occurs at burial depths ranging from 200 m to 9150 m (Mazzullo & Harris 1991). The modern literature on deep-seated carbonate reservoirs provides ample evidence for macroscopic dissolutional porosity formed in situ (e.g., Heward et al. 2000; Korobov & Korobova 2006; Smith 2006, among many others). However, some authors still deny the very possibility of significant dissolution porosity creation in the mesogenetic realm (Ehrenberg et al. 2012). In fact, carbonate reservoir geologists are still largely ignorant of the developments in hypogene karst studies, and stick to the paleokarst concept in interpreting deep-seated solution porosity.
The period since 1990 has witnessed an exponential growth in the number of empirical studies of different kinds of hypogene speleogenesis in various regions around the world. An overview of these works is beyond the scope of this paper. Rich bibliography on hypogene caves can be found in major recent general texts on karst and caves (Ford & Williams 2007; Palmer 2007) and theme-focused monographs and collections of papers (Klimchouk 2007; Klimchouk & Ford 2009; Stafford et al. 2009; Klimchouk at al. 2014). A search in Karst-Base (2015) returns over 360 publications for the “hypogen” keyword, mostly of the last decade.

Several generalizations can be made here about distribution and patterns of hypogene karst porosity (for detailed discussions see Klimchouk 2007, 2012, 2013a, 2013b; Audra 2009). Hypogene speleogenesis occurs in various tectonic and geological/hydrogeological conditions and in rocks of different compositions (all kinds of carbonate rocks, gypsum, conglomerates, sandstones, and quartzites) and ages (from Neoproterozoic to Pleistocene). Its distribution is not limited to continents. With the advent of new sensing technologies, evidence grow rapidly that hypogene karstification occurs in the seafloor (e.g., Betzler et al. 2011; Chen et al. 2015), although its proper interpretation is often hindered due to limitations of the traditional paradigm of karst (Michaud et al. 2005). The depth limit for hypogene speleogenesis is difficult to establish, but available evidence suggest that it occurs at least within several kilometers. It localizes where ascending flow and disequilibrium conditions causing dissolution were supported, continuously or intermittently, during a sufficiently long time, - mainly in zones of discharge and/or interaction of fluid flow systems and regimes of different nature, depth and scales. The localization is controlled by the particularities of regional hydrogeological structure and geodynamic and geomorphic evolution. Hypogene speleogenesis results in a variety of patterns of void-conduit systems, which broadly group into three categories: (1) stratiform, (2) cross-formational, and (3) combined. Patterns and the morphology of hypogene void-conduit system exhibit functional organization that evolved to progressively facilitate ascending flow and discharge. The hydrogeological role of hypogene speleogenesis lies in localized increase of the vertical permeability of separating aquitards, concentration of ascending flow, enhancement of the hydraulic connection of aquifers in layered confined systems and segments in cross-formational fracture-vein systems, and eventually – in improving conditions for ascending discharge.

One of the main reasons for the distinctions in patterns between epigene and hypogene void-conduit systems is the specifics of hypogene speleogenetic mechanism caused by particularities of hydrodynamic behavior of confined flow systems. In confined (semi-confined) settings, in zones where flow is directed transversely upward across layers and formations, both recharge to fractures in soluble rocks and discharge out of them occur through adjacent insoluble beds with a relatively conservative permeability. In cross-formational fracture-vein systems (e.g. fault-controlled), flow crosses soluble and insoluble rocks. Discharge in the whole groundwater flow system is controlled by the least permeable elements in the geological cross-section. Before the onset of speleogenesis, less permeable beds in layered systems are commonly represented by soluble rocks, and discharge through fractures is controlled by their hydraulic capacity. When transverse proto-conduits reach the breakthrough condition, their further growth does not accelerate dramatically, because at some point the control over discharge switches to the permeability of adjacent or more distant insoluble beds, or to unaltered insoluble segments in fracture-vein systems. This switch to the external conservative control over discharge inhypogene speleogenesis subdues the positive feedback loop and the speleogenetic competitiveness and allows adjacent flow pathways to continue their growth, favoring formation of pervasive, maze-like patterns (Klimchouk 2000). This effect has been confirmed by numerical modeling of hypogene speleogenesis in a stratified aquifer system with dispersed basal recharge to the soluble bed (Birk 2002; Birk et al. 2003; Rehrl et al. 2008, 2010). Modeling of conduit development by hydrothermal dissolution along localized cross-formational fractures (Andre & Rajaram 2005; Rajaram et al. 2009) revealed that the thermal coupling between the fluid and rock also causes the suppression of the flow-growth feedback and speleogenetic competition soon after breakthrough. Another specific feature of hypogene speleogenesis is the great role of buoyancy circulation (Klimchouk 1997b, 2007), which has been confirmed and thoroughly studied by thermo-hydrochemical modeling (Chaudhuri et al. 2013).

Dramatic advances in studies of hypogene speleogenesis during last 25 years resulted in that the notion of hypogene karst has changed from an aberrant curious phenomenon to one of the fundamental categories of karst of comparable importance with epigene karst. Recognition of hypogene karst in this capacity clearly signifies an ongoing major shift in karst paradigm, previously overwhelmingly dominated by the epigene concepts and models. Hypogene speleogenesis has broad and important implications for many applied fields such as characterization and modeling of reservoirs in soluble rocks, oil field prospecting and exploitation, geological engineering, mineral resources industries and groundwater management.
The term “endokarst” (endogenous karst) is widely but misleadingly used in the literature to denote underground karst features, and the term “exokarst” (exogenous karst) is used for surficial features. This practice is not consistent with the meaning universally accepted in geosciences, where the term “endogenous” refers to phenomena caused by forces originating from within the Earth, and the term “exogenous” refers to the processes that derive their energy from external sources. To be compatible with this usage, the term “exokarst” should be used largely in a sense of the notion of epigene karst, but also include artesian speleogenesis in the upper hydrodynamic storey of basins, driven by meteoric topography-controlled circulation with no involvement of deep endogenous flow systems. Karstification, produced by fluid flow systems driven by internal sources of energy, even those containing waters of the meteoric origin, can be regarded as a realm of “endogenous karst”. Thus, this meaning of endogenous karst is close to the notion of hypogene karst, but it is not entirely equivalent to it as the latter also includes speleogenesis driven by topography-controlled artesian circulation.

An original concept of endokarst has been suggested by Ezhov & Lysenin (1990) and further developed in Ezhov et al. (1992). According to these views, the realm of endokarst encompasses the parts of the crust below so called “buffer” zone – a dense interval of maximum compaction of the rock and complete closure of all types of porosity at depth of about 7-15 km, which forms a planetary-scale regulator of defluidization of the deeper parts of the Earth’s crust. Endokarst processes involve liquid-vapor fluids, released from thermal breakdown of hydrous minerals and arriving from the lower crust and the upper mantle, acting at temperatures above 100 °C and pressures approaching the lithostatic ones (Fig. 1). In such conditions the fluids are highly aggressive with respect to many sedimentary, metamorphic, and igneous rocks. Fluid-filled porosity may exist in the endokarst story because fluids are under lithostatic pressures so that pressure gradients are negligible. However, cavities can be preserved while passing through the above buffer zone only if filled with some secondary mineral such as calcite or anhydrite.

Although dissolution processes certainly operate in the zone of lithostatic pressures, little is known about conditions at such depths (geological inhomogeneities, dynamics of fluid and their physico-chemical parameters) which may control concentration of fluid flow and dissolution effects, i.e. formation of void-conduit systems. It is obvious that the mechanisms of speleogenesis known to operate in the upper parts of the Earth’s crust (hydrostatic zone), are not applicable for the lithostatic zone. As speleogenesis is the most essential attribute of karst, it is questionable whether the phenomena hypothesized by Ezhov et al. (1992) can be classified as karst.

It is much more certain that pulse breakthroughs of deep fluids into the hydrostatic zone, that propagate upward along deep-rooted faults and other heterogeneities, transecting this zone to various heights and up to the surface, play the important role in generating hypogene speleogenesis. Such pulses interact with various ground-

---

**Fig. 1:** Vertical hydrodynamic zoning (A) and karst stories (D) of the Earth’s crust. (B) and (C) show, respectively, dominating flow regimes and the origins of groundwater in different zones and stories. (A) also shows changes of some important parameters with depth: solid line – parameters before the impulse breakthrough of fluids through buffer subzone II-A; dotted line – parameters after the breakthrough. (A) is from Ezhov & Lysenin (1990), as reproduced in Andreychouk et al. (2009).
water flow systems and are believed to be responsible for pressure anomalies and associated thermal, hydrochemical, and gas anomalies. For more details and discussions of the endokarst concept of Ezhov and Lysenin see Andreychouk et al. (2009) and Klimchouk (2012).

GENETIC TYPES OF KARST

The postulates about the primary importance of groundwater flow in speleogenesis, and of speleogenesis in the formation of karst, substantiate a proposition that genetic types of karst are to be distinguished based on types of speleogenesis.

Two fundamental types of speleogenesis, hypogene and epigene, are determined mainly by distinct hydrodynamic characteristics of the respective groundwater flow systems: (1) stratiform confined aquifer systems, or cross-formational fracture-vein systems, of varying depths and degrees of confinement, and (2) hydraulically open, near-surface unconfined systems. Accordingly, two major genetic types of karst are distinguished within the upper part of the Earth's crust: hypogene karst and epigene karst. They differentiate due to fundamental differences in boundary conditions, lithological, structural and geochemical conditions and hydrodynamic regimes of groundwater (fluid) flow and speleogenesis (Fig. 2), as well as due to differences in the evolutionary trajectories of corresponding karst systems.

Epigene and hypogene types of karst differ in many characteristics, notably in relationships with the surface, hydrogeological behaviour, groundwater quality, and economic resources they may contain. This determines substantial differences in their environmental impacts, the areas of practical importance and approaches to solving karst-related issues.

Flank-margin speleogenesis (Mylroie & Carew 1995) is often distinguished as a particular type, based on singularities of cave development caused by high matrix porosity of young carbonates and localization of dissolution in the freshwater/seawater mixing zone, particularly in the margins of coastal freshwater aquifers / lenses in islands. This can be taken as a ground for distinguishing a particular genetic type of coastal / eogenetic karst. The standard flank-margin speleogenetic model (Mylroie & Carew 1995) was based on an assumption that the rock sequence is homogenous, and the freshwater lens was considered as an unconfined aquifer. However, taking into account that considerable layered heterogeneity and the leaky confinement can be present even in young carbonates, speleogenesis in the margins of freshwater lenses can be caused by flow rising across less-permeable beds, i.e. it can be truly hypogene in the hydrogeological sense (Fig. 3 B; Klimchouk 2014).

Fig. 2: Karst and speleogenesis in the context of diagenetic zones and groundwater flow regimes. The diagram is out of scale and the vertical dimension is greatly exaggerated. 1 - meteoric, topography-driven regime: a - local systems (unconfined), b - regional and sub-regional systems (confined); 2 - expulsion (exfiltration, basal) regime, commonly overpressured, driven by compaction and tectonic compression: a - in newly-deposited sediments, b - in older rocks; 3 - interfaces between groundwater regimes and systems: a - meteoric/expulsion regimes, b - local/regional-subregional meteoric systems; 4 - poorly permeable beds (only a few are shown on the diagram); 5 - meteoric flow paths; 6 - basal flow paths; 7 - enhanced cross-formational communication; 8 - intense gas inputs; 9 - temperature and gradient anomaly: positive, negative; 10 - redox conditions: oxidizing, reducing; 11 - epigene speleogenesis; 12 - hypogene speleogenesis. From Klimchouk (2012).
EVOLUTIONARY TYPES OF KARST

It was implied for a long time that karst development commences only with the exposure of a soluble formation to the surface. Within the traditional paradigm of karst, its evolution has been viewed mainly from the perspective of the geomorphological evolution. With the establishment of the concepts of interstratal karst (Quinlan 1978), deep-seated karst and hypogene karst, it became obvious that evolution of karst should be viewed from the perspective, and in timescales, of the entire life of a geological formation, in which the geomorphogenesis is commonly the latest stage in a sequence of others. In sedimentology, environments of alteration of sedimentary formations are treated in terms of eogenesis, mesogenesis and telogenesis (Choquette & Pray 1970), the successive stages in the normal cycle. Since the hydrogeological context is the most important for karst, its evolution should be viewed as a part of the evolution of the water-rock system in response to diageneric and tectonic processes in the course of burial, uplift, denudation, and geomorphic development.

A useful framework to characterize the changes in major characteristics of karst is provided by the classification of karst settings in the context of the geological evolution of a soluble formation (Fig. 4). This classification was developed by Klimchouk (1996) and Klimchouk & Ford (2000) in the form of an evolutionary scheme, using earlier ideas and terminology by Ivanov (1956) and Quinlan (1978). Different types of karst (settings) represent the potentially successive stages (states) of its evolution, between which the major boundary conditions (e.g. recharge/discharge), the overall flow patterns and regimes, and extrinsic factors and intrinsic mechanisms of speleogenesis change considerably.

These types are (in the order they potentially evolve) syngenetic/eogenetic karst in freshly deposited rocks; deep-seated karst, which develops during mesogenesis, particularly during its ascending limb (when the rocks are being shifted toward the surface); subjacent karst, where the cover is locally breached by erosion and direct hydraulic interaction with the surface is established; entrenched karst, in which valleys incise below the bottom of the karst aquifer and drain it, but where the soluble rocks are still covered by insoluble formations for the most part; and denuded karst, where the insoluble cover materials have been completely removed. If the soluble rock bypassed burial, or karstification commenced solely after the rock was exposed after burial, such karst represents the open karst type. Deep-seated karst, subjacent karst, and entrenched karst represent the group of intrastatal karst types, whereas denuded and open karst form the group of exposed karst types. Later on, karst may become mantled by a cover that develops contemporaneously with the karst (mantled karst), or reburied under younger rocks to form paleokarst, and be re-exposed (exhumed karst).

Although this classification does not directly specify the origin of caves, it characterizes dominant speleogenetic modes in different environments. The evolutionary types of karst correlate with types of speleogenesis (genetic types of karst) in the following way. Syngenetic/eogenetic karst domain has some particu-
Karst systems receive the expression in the landscape and directly interact with external landscape-forming factors, and themselves become a factor of geomorphogenesis, only at certain stages of the development, when the soluble formation is originally exposed to the surface (syngenetic karst type), or is transferred into the shallow subsurface after burial in the course of uplift and denudation (entrenched and denuded types of karst). At the stage of deep-seated karst, hypogene karst systems commonly have no geomorphic expression.

Deep-seated karst is represented exclusively by hypogene speleogenesis. In subjacent karst settings both hypogene and epigene speleogenesis may operate, depending on a dominant groundwater regime and interaction between different flow systems, but hypogene speleogenesis often dominates. Entrenched and denuded karst types are overwhelmingly epigenic, with inherited hypogenic features that can be reworked by epigenic processes or get fossilized. In both karst types, however, ceasing hypogene systems may still operate. Open karst is marked by exclusively epigene speleogenesis.

Hypogene and epigene types of karst are characterized by fundamentally different relationships with geomorphogenesis and landforms. Epigene karstification occurs in the near-surface conditions and is directly linked with recharge from the immediately overlying or adjacent surface. Accordingly, it is directly linked with the landscape. It is subordinated to the gross landform development that creates the particular configuration of exposure, recharge and drainage for the soluble formation, i.e. the initial pattern of hydraulic gradients and

RELATIONSHIPS BETWEEN KARST AND GEOMORPHOGENESIS

Fig. 4: Evolutionary types of karst and speleogenetic environments (modified from Klimchouk & Ford 2000). Background colors indicate the domains of hypogene and epigene speleogenesis.
groundwater flow. Landscape is one of the determining factors in the early epigene karstification. Epigene karst features are roughly coeval or younger with respect to major landforms. In mature stages, epigene karstification itself becomes the important factor of geomorphogenesis at the meso-scale, as in tower- and cone karst landscapes.

Hypogene karstification is not connected with the local surface recharge, being driven by the ascending flows between aquifers in confined aquifer systems and along cross-formational fault/fracture zones. Landscape features at mega- and macro-scales indirectly affect hypogene speleogenesis within the first, and sometimes, the second hydrogeological storeys in large cratonic basins, as they determine the pattern and intensity of transverse flows between stratal aquifers. Hypogenic speleogenesis localized along cross-formational disruptions may not be related to the landscape at all. In the context of the long-term geologic and geomorphic development, geomorphogenesis indirectly affects hypogenic karstification through changes in the boundary conditions of confined aquifer systems on their upper contours, i.e. through erosional dissection and denudation of the upper confining unit.

Recognition of the possibility that hypogene void-conduit systems can develop at depth, largely independent of the surface, leads to revisiting general ideas about the relationship of karst and geomorphogenesis. Hypogenic karst systems can be significantly older than the modern landscape. When a hypogenically karstified formation is brought to the shallow subsurface by uplift and denudation, the karst system interacts with geomorphogenesis through a different scheme than in the case of the epigene karst. Its interaction with the landscape includes: focusing ascending groundwater discharge (with respective contribution to localization and development of fluvial erosion features), collapsing of large cavities, intercepting surface runoff and focusing it along unroofed conduits, vertically enhanced disintegration of rock masses along rift-like conduits and formation of cliffs and outliers, exposing unusual relict karst morphology in cliffs, etc. Thus, karstification is not subordinated to the overall relief development as in the case of epigene karst, but geomorphogenesis at a certain stage can be largely controlled by intercepted hypogene karst structures, as shown recently for the Crimean Piedmont (Klimchouk et al. 2013a, 2013b). I strongly suspect that many unusual cliff-, canyon-, butte- and pillar-dominated landscapes in carbonates and sandstones (such as, for instance, Meteora in Greece, Petra in Jordan, and Poseidon system in the Bohemian massif, Czech Republic, among others) could owe their origin to the disintegration of hypogene rift-dominated conduit systems, although special studies are needed to demonstrate this.

### CLARIFYING THE NOTION OF KARST

Several trends are apparent in modern karstology that take the notion of karst well beyond the limits of the traditional, largely geomorphological, paradigm of karst:

- Acknowledgment of the central role of speleogenesis in the formation of karst;
- Recognition of the primacy of the fluid flow in the development of karst (i.e. of the hydrogeological essence of karst);
- Recognition of the wide occurrence and peculiar characteristics of hypogene karst;
- Adoption of the broad perspective to karst evolution which goes beyond the contemporary geomorphological epoch and encompasses the entire life of a geological formation.

These developments are changing views on which properties of karst are essential.

**WHICH ATTRIBUTES OF KARST ARE ESSENTIAL?**

Based on the overview provided in Section Variants of definitions of karst and subsequent discussions, it is briefly examined below if the attributes of karst used in the definitions are really essential. It is accepted here that essential properties are those that the object must have (i.e. necessary properties). It is also important to identify properties, or a combination of properties, that make the object unique (i.e. exclusive properties).

- **Dissolution.** It is universally accepted that dissolution is an essential process in karstification, and references to dissolution are included to most of the definitions of karst. Moreover, many definitions literally state that karst is the result of dissolution of rocks [1, 6, 8, 9, 15].

There are two problems here. One is that dissolution is ubiquitous in the Earth’s crust, and it is not exclusively attributed to karst. Referring to this property alone is not sufficient to define karst. Another problem is that karst (karstification) is commonly identified with dissolution, – a source of a widespread misunderstanding in the geological literature. For instance, in the literature on carbonate reservoirs the term “karst” is rarely
used even with regard to macro-scale solution porosity and permeability features; the notion of karst is effective-ly substituted by the notion of dissolution. The practice of equating karst to dissolution is also traceable through the karst literature.

This is an obvious case of misleading reductionism. Karstification is not equal to dissolution. Dissolution is a chemical process, whereas karstification (including speleogenesis) is a hydrogeological mass transfer / mass transport process, in which fluid flow and chemical dissol-ution are coupled, and the process is governed by the evolution of the geological environment. Adding to the complexity is that, whereas dissolutional removal is certainly one of the essential attributes of karst, other de-structive processes also take part in karstification, and their “weight” in the overall process increases with the maturation of the karst systems.

Carbonate rocks, soluble rocks. It was shown in Sec- tion Phenomena in carbonate rocks that the restriction of the notion of karst by relevance to only carbonate rocks is too specific, misleading, and outdated. Defining karst as phenomena in soluble rocks is, contrarily, too vague; – the nature of the phenomenon that makes it distinct from other phenomena is still to be additionally speci-fied. The qualification of rocks as (readily) soluble ones was originally set in parameters of the near-surface con-ditions. The solubility of a rock depends on the physical and chemical properties of the solute and solvent, as well as on temperature, pressure and the pH of the solution, and extension of the karst domain to hypogene environ-ments makes these properties varying in a much wider range than it was thought earlier. Accordingly, the list of rocks that can be deemed as easily soluble and potentially karstifiable, is being expanded. For instance, solubility of quartz at temperatures of 300–350 °C and pressures of 200–250 MPa becomes comparable to that of gypsum or anhydrite in near-surface conditions (Ezhov et al. 1992). Some large caves in quartzites have been recently shown to be of deep-seated hypogenic origin (Sauro et al. 2014).

Landscape characteristics, specific landforms. As shown in Section Karst as a specific topography, land-scape, or a set of landforms (physiognomic approach)and elsewhere in this account, the attribute of specific char-ac-teristics of landscape, or even of the very presence of land-scape, are not essential for karst. Wide usage of references to landscape/landform characteristics in definitions of karst is the result of inertia from the previous paradigm. During the last 25-30 years the notion of karst has clearly decoupled from the surface-related attributes that is one of indications of the paradigm shift in karstology.

Water (fluid) – rock interaction. The notion of a fluid-rock interaction, used in some process-based defini-tions [17, 21], does potentially encompass the coupling of fluid flow and chemical dissolution, and is therefore an adequate and essential designation of the main drive for karstification. It is, however, too broad and needs in additional indications to the nature (mechanism) of the interaction. Ezhov et al. (1992; [18]) specified it as a metasomatic (alteration) process, based on the concept of Pospelov (1973) about “extended” metasomatism as a “throughgoing” (overarching) process in which opera-tion of the zones of dissolution, transport, and deposition is “stretched” in time or space or both.

Fluid circulation aspects. The paramount importance of groundwater circulation in karst was recognized by many scholars. For instance, Sweeting (1972, p. 5) noted that “... the sinking of water and its circulation under-ground is the essence of the karst process ...”. The attribute of the groundwater circulation (movement) is mentioned in some form in six definitions of the selection used in this paper [4, 11, 12, 16, 17, 19], but it is clearly empha-sized only in two of them, – in those by Devdariani (1962, [19]) and Huntoon (1995, [12]). As underscored in Section The supremacy of groundwater flow in speleogenesis and karst development, the character of the water-rock interaction in a given rock medium and the distribution of the effects are determined by the nature, intensity and pattern of the groundwater circulation, i.e. by hydrodynam-ic characteristics of groundwater flow systems. The attribute of the groundwater circulation is therefore the most essential for defining karst.

Self-organization. Only two definitions include this attribute, again by Devdariani ([19]) and Huntoon [12]. Since self-organization of permeability and groundwater flow in soluble rocks is the essence and the main result of speleogenesis (see sections Speleogenesis as a primary process in the formation of karst and Self-organization of permeability and flow systems), and since karst is a function of speleogenesis (see Section Speleogenesis as a pri-mary process in the formation of karst), the attribute of self-organization should be regarded as one of the most essential for karst. Huntoon (1995) and Worthington & Ford (2009) emphasized the role of self-organization of conduit permeability in formation of karst aquifers, and Klimchouk (2011) considered it to be a system-form-ing property of the karst ecosystem. One of the results of self-organization of the flow system is the dynamic and dramatic increase in the aquifer heterogeneity and permeability (i.e. the efficiency of flow), which can be deemed as the unique characteristic of karst.

Concentration of flow. The definition of Devdariani [19] is the only one which includes the attribute of flow concentration in the definition of karst. It is a direct con-sequence of self-organization of permeability and flow in soluble rocks (i.e. speleogenesis), and one of the ma-jor characteristics of mature karst systems (Worthington
et al. 2000; Worthington & Ford 2009). It is therefore an essential attribute of karst.

Localized occurrence. This is a typical feature of distribution of karstic voids and conduits, the result of concentration of flow and increased heterogeneity of the karstified media. Tsykin (1985) considered the localized occurrence as an invariant property of karst. This attribute is tacitly implied by the notion of conduit porosity.

Presence of cavities. Cavities as an attribute of karst are mentioned in several definitions, together with landscape features. Unlike surface landforms, the presence of solutionally enlarged cavities (void-conduit networks) is the inherent, and hence essential, attribute of karst systems.

Transformation and destruction of rocks. Since dissolution is a form of destruction of rocks, most definitions imply destruction as they mention dissolution. Sokolov (1962; [16]) defined karst as a process of destruction and obliteration of rocks, which means destruction in the geological sense. Tsykin (1985; [14]) defined karst as a totality of forms of selective destruction of rocks. Ezhov et al. (1992; [18]) presented karst as a process of metasomatic alteration of rocks. Transformation and destruction of rocks are indeed the essential attributes of karst.

In summary: The attribute of peculiar landscape/set of landforms, which is most widely used in definitions of karst, is in fact not essential in the light of the modern understanding of the phenomenon. Other properties considered above are essential, but only one of them is an exclusive one for karst, namely the dynamic and dramatic increase in the heterogeneity and permeability of the media, the result of self-organization of the flow system and speleogenesis. It is possible to determine a combination of the essential attributes which in aggregate would uniquely define karst and state its essence.

REFINING THE DEFINITION OF KARST

In existing definitions of karst different categories have been used for the object itself: a landscape (a set of landforms), a terrain (a territory), a geological environment, a totality of forms, a process, and a system.

It is suggested here that the system-based approach to the notion of karst holds the best promise to grasp and adequately represent the specific nature of the karst phenomenon. It is also suggested that this notion should rely on the concept of a groundwater (fluid) flow system. The karst system can be uniquely defined only in terms of groundwater flow.

The presence of soluble rocks in the geological environment causes the phenomenon of self-organization of the permeability and flow pattern (i.e. speleogenesis) which determines the specific evolution of the groundwater flow system and transforms it into a new quality (state), – karstic. Therefore, karst can be viewed as a specific groundwater (fluid) flow system, peculiar properties of which have developed as the result of speleogenesis. A similar approach has been already applied with regard to karst aquifers (Worthington & Ford 2009). It should be expanded to karst as a whole, since almost all attributes deemed to be essential for karst are the results of the specific (i.e. speleogenetic) evolution of the groundwater flow system in soluble rocks.

Of the definitions considered in this paper, the one by Huntoon (1995; [12]) encompasses, explicitly or tacitly, the most of essential properties attributed to the phenomenon. The pattern of the Huntoon formulation is used here to refine a definition of the notion of karst in the light of the above discussion, with some minor modifications that follow from the above discussions.

Karst is a fluid flow system (geohydrodynamic system) with a permeability structure evolved as a consequence of dissolutional enlargement of initial preferential flow pathways, dominated by interconnected voids and conduits, and organized to facilitate the circulation of fluid in the downdip direction due to the positive feedback between flow and conduit growth.

This approach implies the use of a subordinated system of key concepts and terms, which revisited definitions are suggested below.

Speleogenesis (karstic) – the formation of voids and conduits in rocks through mainly dissolutional enlargement of initial preferential flow pathways involving self-organization due the positive feedback between flow and conduit growth.

Karst (karstic) process (syn. karstification, karstogenesis) – a geological process (an interconnected set of processes) of transformation of soluble rocks under the dominant action of coupled flow-dissolution processes and respective self-organization of the groundwater flow system. Karstification is manifested in the emergence, development and degradation of speleogenetic (macroscopic) porosity, in increasing permeability, enhancing heterogeneity and anisotropy of hydraulic, reservoir and mechanical properties of rocks, as well as in the respective evolution of the karst geohydrodynamic system.

The progressive evolution of the karst system is dominated by processes of solutional removal of matter from the host rock (speleogenesis), with increasing intensity and concentration of fluid circulation, and heterogeneity. The regressive evolution is characterized by predominance of various paragenetic processes of precipitation/sediment accumulation, fossilization and disintegration of void-conduit porosity structures.
The karst process at some stage may include the formation of specific karst topography and landscapes as a result of interaction of the karst geohydrodynamic system and underground karstic features with the surface and external agencies.

**REFERENCES**


**CONCLUSIONS**

The notion of karst within the traditional paradigm is largely geomorphologic (physiognomic), build around external appearance of a karst system developed in exposed and shallow-lying soluble rocks under the direct action of surface recharge. This notion corresponds to what is now termed epigenic karst.

General understanding of karst has changed during last 30-40 years. Dramatic advances in studies of karst hydrogeology and speleogenesis have revealed that the essence of karst lies not in its morphological characteristics but in the evolution of a groundwater flow system in soluble rocks, and that speleogenesis is the principal intrinsic mechanism of this evolution. The progress in understanding of hypogenic karst, stimulated by speleogenetic researches and massive new data arising from hydrocarbon prospecting and exploration, have led to its recognition as another fundamental type of karst. Besides gravity-driven flow systems, hypogenic karst can be related to systems driven by endogenous energy sources, and its evolution may go far beyond the current geomorphologic epoch.

With these developments, views on essential attributes of karst have been clearly changing, indicating the ongoing shift from the largely geomorphologic paradigm of karstology to the geological (hydrogeological) one. This shift, however, is not adequately reflected in definitions of the notion which are in a broad use in the earth-science literature. A refined approach is suggested where karst is viewed as a specific fluid flow system (geohydrodynamic system), which has acquired its peculiar properties in the course of speleogenesis. Underground and superficial karstic features are the reflection of the functioning of the karst geohydrodynamic system in the present (active karst) or in the past (relict karst and paleokarst).

**ACKNOWLEDGEMENTS**

The author thanks to the reviewers, one of which was Prof. Derek Ford, for constructive comments and suggestions, which considerably improved this paper.


Birk, S. 2002: Characterisation of karst systems by simulating aquifer genesis and spring responses: Model development and application to gypsum karst.- Tübinger Geowissenschaftliche Arbeiten C60, pp. 122, Tübingen.


Dublyansky, V.N., 1980: Hydrothermal karst in Alpine folded belt of southern part of the USSR. - Kras i speleologija, 3, 18–33.


Esteban, M., & J.L.Wilson, 1993: Introduction to karst system and paleokarst reservoirs. - In: Fritza et al. (eds.) Paleokarst Related Hydrocarbon Reservoirs. SEMP Core workshop no. 18, pp. 1–9.


Klimchouk, A.B., 1992: Large gypsum caves in the Western Ukraine and their genesis.- Cave Science, 19, 1, 3–11.


Klimchouk, A.B., 2014: The methodological strength of the hydrogeological approach to distinguishing hypogene speleogenesis.- In: Klimchouk, A. et al. (eds.) Hypogene Cave Morphologies, Karst Waters Institute, pp. 4–12, Leesburg, Virginia.


Quinlan, J.F., 1978: Types of karst, with emphasis on cover beds in their classification and development.- PhD thesis, Univ. of Texas at Austin, pp. 323.


