1

Scientific and geological context

CHAPTER 1
Introduction to speleothems and systems

For paleoclimate, the past two decades have been the age of the ice core. The next two may be the age of the speleothem.

Gideon Henderson (*Science*, 2006)

1.1 What is all the fuss about?

Moore (1952) recognized a need to specify an unambiguous term for mineral deposits that grew within caves and proposed 'speleothem' (Greek: *spelaion*, cave; *thema*, deposit). In recent years speleothems have been established as one of the most valuable resources for understanding Earth surface conditions in the past, from times when glaciers waxed and waned and our human ancestors emerged, to the present day. By 'conditions' we mean not only the local context (soil, vegetation, landscape instability and climate), but also the regional to global patterns of change that characterize former environments and climates (Fig. 1.1).

Because no two speleothems are identical (Plate 1.1), they were formerly thought to be too complex to generate reliable archives of the past. It has taken a considerable effort over the past 40 years both to show that reliable records can be obtained which, in some cases, display global phenomena, and to demonstrate that speleothems form by a set of processes that can be rationalized and understood. Even in the late 1990s, textbooks on palaeoclimates and palaeoenvironments treated speleothems only briefly. Now it can be argued that certain long and well-dated records provide the most definitive archives of the global environmental system. Speleothems also provide an enormous resource for future research on past changes, with an ultimate dynamic range of eight orders of magnitude from days to a million or more years.

The rest of section 1.1 provides a summary of the essentials of speleothem science for the benefit of all those, from undergraduates to specialists in related fields, who want to get to first base. In section 1.2, we explain how the rest of the book is organized.

1.1.1 What types of speleothem are useful for generating climate archives?

Many different forms of speleothem can be distinguished (Hill & Forti, 1997). Here we consider primarily two types of deposit (*dripstones*) that grow from dripping water: *stalactites* (Greek: *stalaktós*, dripping) growing down from the cave roof and *stalagmites* (Greek: *stalagmós*, dropping) building up from the cave floor. Also pertinent are more continuous deposits (*flowstones*) that accrete beneath thin sheets of water on cave walls and floors (Fig. 1.2a). Finally, we consider some long records obtained from crystalline deposits that form underwater. Flowstones tend to have fairly continuous layers and so it is possible to duplicate records by sampling (e.g. by coring) in different places.

Speleothem Science: From Process to Past Environments, First Edition. Ian J. Fairchild, Andy Baker.

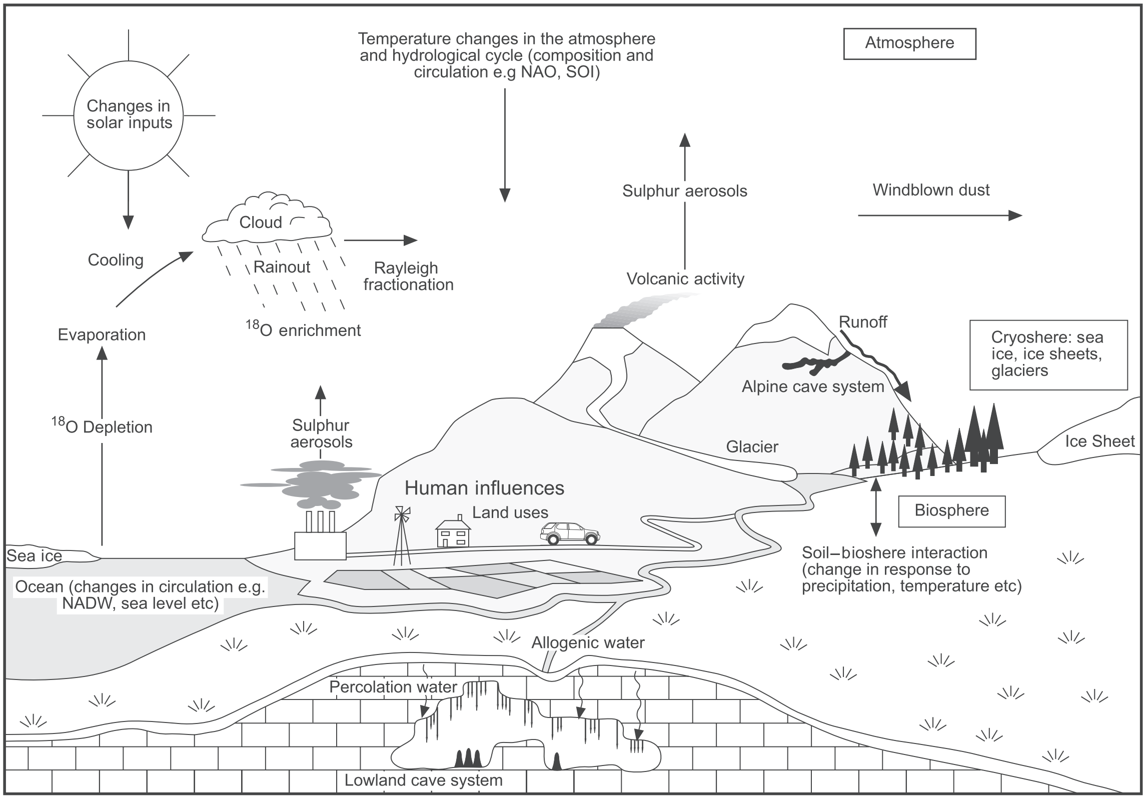

Fig. 1.1 Speleothems as underground recorders of signals related to parameters of the external Earth system (Fairchild et al., 2006a).

However, the layers can show a small-scale topography reflecting the ponding of surface water (Fig. 1.2a). Stalagmites are more commonly used to generate archives than stalactites because their internal structure is simpler. Fig. 1.2b illustrates a stalagmite cross-section illustrating that there is internal layering which tends to be flat on the top of the sample, allowing a set of observations representing different time periods in the past (*time series*) to be generated along a sub-vertical line. However, the sample illustrated in Fig. 1.2b displays lateral shifts in its growth axis, related to changes in the landing position of water drops. It is also commonly found that growth may pause for an extended period and then resume, and that these characteristics reflect either changes in climate or local processes within the aquifer in different cases.

1.1.2 Where do speleothems occur?

The speleothems used to find out about the past are almost all *calcareous* (made largely of calcium carbonate, $CaCO_3$) and composed of the minerals calcite and/or aragonite. Such speleothems occur within *carbonate rocks*, typically *limestones* ($CaCO_3$) and/or *dolomite* ($CaMg(CO_3)_2$. These rocks store, transmit and yield water readily and hence are *aquifers*. The *permeability* of aquifers refers to the ease with which they transmit water. Carbonate aquifers occur in *karstic* regions (Gunn, 2004; Ford & Williams, 2007) where there is little surface water because it drains readily into the bedrock (Fig. 1.3). The cavities in the bedrock typically range from tiny pores, through enlarged fissures to *conduits* which used to host, or still contain, underground streams. Over time, the drainage system

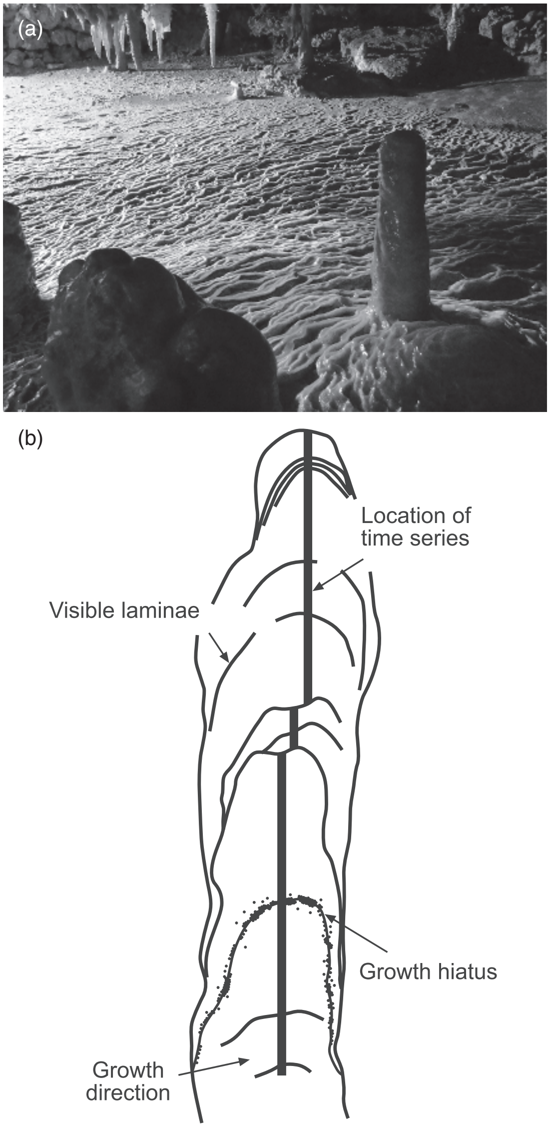

Fig. 1.2 (a) Speleothems in Grand Roc Cave, Dordogne, France. The cave floor is covered with a flowstone on which have developed walls (terracettes) and intervening gour pools. A columnar stalagmite a few centimetres wide occurs to lower right and stalactites in the upper left. (b) Cross-section through a stalagmite illustrating visible growth layers (laminae) and the shifting position of the growth axis which represents the ideal position to generate the most continuous time series.

develops and tends to lower the depth of the *water table*, the regional surface below which cavities are water-filled. The water table divides the aquifer into the underlying *phreatic* and overlying *vadose* zones, although in karst terrains the division is not always simple. Pragmatically, in this book we refer to the vadose zone, from which dripwater originates, as part of the aquifer, although conventionally the term is often restricted to the phreatic portion of the rock. Caves in vadose zones tend to cut vertically downwards, whereas in the phreatic zone they develop as passages elongated in the direction of water movement. Speleothems are normally formed in abandoned passages (Fig. 1.3) in the vadose zone and are fed by water passing through the soil into the uppermost karst, which is typically a zone of significant water storage (the *epikarst*). Speleothems form part of a long and complex history of cave spaces. Hence they can be interlayered with particulate sediments and are ultimately may be later broken by earthquakes or human disturbance, buried, flooded, eroded or dissolved.

Modern calcareous speleothems occur in nearly all continental regions and are principally limited by the availability of karstic host rocks and liquid water. However, they tend to be only weakly developed in frigid regions where there is less chemical drive for them to form (Fig. 1.4b; section 1.1.3). They are found in many modern desert regions, having formed during more humid conditions in past millennia. Speleothems that originally formed not far above the water table can be found on mountain tops in tectonically uplifting regions. Conversely, precipitates that formed during previous ice ages when sea level was much lower are observed beneath the sea and can be recovered for study by divers.

1.1.3 How do they form?

The essential chemical processes that lead to speleothem formation are illustrated in Fig. 1.4a and described in its legend. Essentially there is a carbonate dissolution region in the soil and upper epikarst and an underlying vadose region of $CaCO_3$ precipitation. These correspond to $CaCO_3$-*undersaturated* and -*oversaturated* solutions respectively. The amount

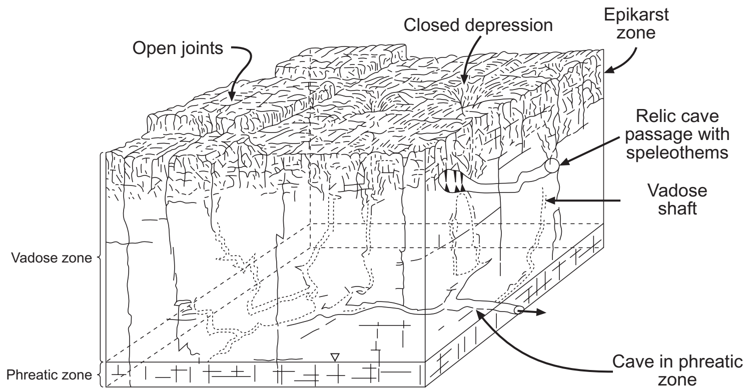

Fig. 1.3 The complex structure of an active karst aquifer in which speleothems typically occur in caves that are no longer being developed and are well above the water table (modified from Smart and Whitaker, 1991).

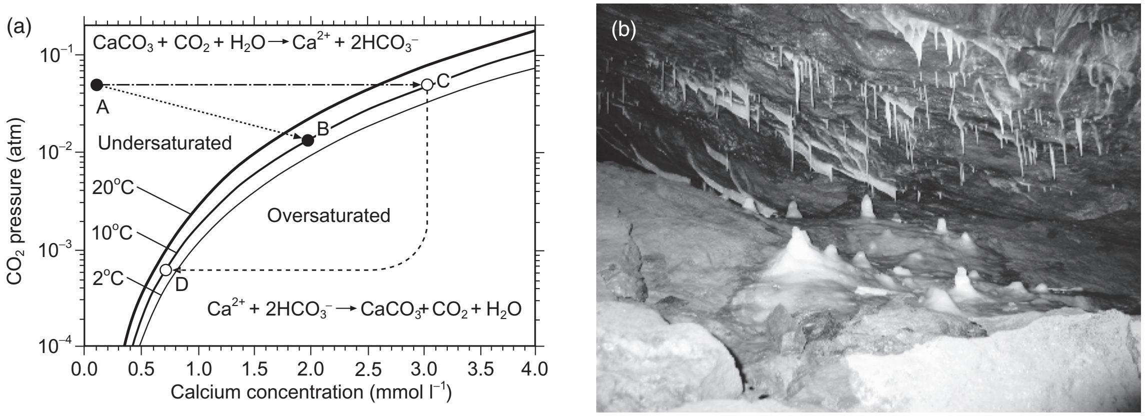

Fig. 1.4 (a) The normal chemical pathway leading to speleothem formation. High carbon dioxide partial pressures (P_{CO_2}) arise in the soil (e.g. point A) owing to respiration and decomposition of organic matter. When percolating water with a high P_{CO_2} reaches carbonate minerals, they will be dissolved, increasing the calcium concentration in solution. If there is no renewal of CO_2, the water follows a 'closed system' path to saturation point B, whereas if CO_2 is replenished to maintain a constant P_{CO_2} then saturation will be reached at point C (both B and C are specific to 10 °C as shown). As the water descends the karst system, at some point it may encounter an air space with a lower P_{CO_2} than the original soil. The water degasses CO_2 and enters the oversaturated field and tends to precipitate $CaCO_3$ (e.g. dashed line C–D). Modified from Kaufmann (2003). (b) Holocene speleothems in a cave above the Arctic circle, Norway: slow growth limited by low P_{CO_2} in the soil zone.

of calcite and dolomite that can be dissolved by water depends on its acidity, normally expressed as the partial pressure of carbon dioxide (P_{CO_2}) with which the water is stable. Although strong acid (e.g. from oxidation of sulphide minerals such as *pyrite*, FeS_2) can be important locally, normally dissolution is enhanced mainly by high P_{CO_2} values of up to several per cent (0.01–0.1 atm) generated by respiration and organic decomposition in soils. As the water descends through the karst, it ultimately encounters a gas phase with a lower P_{CO_2} compared with that which it has previously encountered. This causes degassing of CO_2 from the solution and precipitation of $CaCO_3$. Figure 1.4 shows that the difference in P_{CO_2} between soil and cave is a key control on the quantity of calcium removal from the water, and hence the rate of calcium carbonate precipitation. Growth rate, i.e. upward extension rate in the case of stalagmites, can be of the order of a millimetre per year in humid, warm regions, but is more typically less than 100 μm per year in cool temperate regions. Precipitation can happen above the site where the observer is based and such *prior calcite precipitation* leads to reduced oversaturation and can be identified by a characteristic evolution of water composition. Long-term continuous growth as slow as 1 mm per thousand years has been documented from weakly oversaturated waters.

Speleothem growth requires quite specific conditions of availability of water to supply the ingredients of growth and circulation of air to take away the waste product carbon dioxide. We use the metaphor of the *speleothem incubator* in this book to describe this life-support system that maintains speleothem growth. Although conditions in cave interiors are much more constant than above ground, they are not completely static. Temperature varies near entrances, the humidity and carbon dioxide content of the air change laterally and through the year, and the quantity of infiltrating water is a function of the passing seasons. These changing conditions normally impart an annual visible or chemical lamination within a speleothem if it grows quickly enough to be resolved. This lamination contains information about the seasonality of the climate during deposition. The commonest type is an *annual couplet* (Fig. 1.5a), reflecting warmer–cooler or more usually wetter–drier alternations, but a discrete thin *impulse lamina* (Fig. 1.5b), characteristic of seasonal influx of soil-derived material, is common in cool temperate climates.

1.1.4 How do we date them?

A time series of observations is of limited use unless we can assign real ages to it. In the case of modern, actively accumulating speleothems, their rate of growth can be determined by direct observation, e.g. on human artefacts (Fig. 1.5a). Another successful technique on these materials is to demonstrate a distinct signal within the stalagmite of enhanced levels of radiocarbon (^{14}C) resulting from atmospheric nuclear tests of the 1960s: this has been done for the sample illustrated in Fig. 1.5b (Smith et al., 2009). In both cases, these observations show that the growth lamination displayed is annual and so the duration of growth of older speleothems formed in the same setting can be derived by counting laminae. However, it would be unwise to rely solely on such a method because hiatuses in growth may not have been detected, and in any case growth may not have continued up to the present day. Hence normally an *absolute* dating method is used which allows assignment of a speleothem layer to a particular calendar age. The most common methods are *radiometric*, that is they rely on decay of a radioactive species from a defined starting point. By far the most commonly used method for samples between a few hundred and a few hundred thousand years in age is the uranium-thorium disequilibrium method, but a variety of radiometric and other methods can be used as a check or to extend the dating range to millions of years.

The development of techniques for precise and accurate U–Th dating on small amounts of sample is the single most important factor that has allowed speleothem science to become so prominent in recent years (Edwards et al., 1987; Hoffmann et al., 2007). Work that used to be laboriously undertaken on tens of grams of sample in the 1980s can now be performed on milligrams following the successive development of thermal ionization mass spectrometry (TIMS) and multi-collector inductively coupled plasma mass-spectrometry (MC-ICPMS). The core principle is that as the speleothem grows, it incorporates some uranium from aqueous solution, but fails to incorporate the insoluble element thorium. The nuclide ^{230}Th accumulates over time, by alpha-decay from ^{234}U, after that particular speleothem growth layer has been deposited. The *half-life* (time taken for half the radioactive nuclide to decay) is around 245,000 years and the method can be used to date samples up to around 500,000 years in age. Cheng et al. (2009b) achieved astonishingly good analytical precision (2σ) of 100 years or less on samples over 120,000 years old, which represents the state-of-the-art.

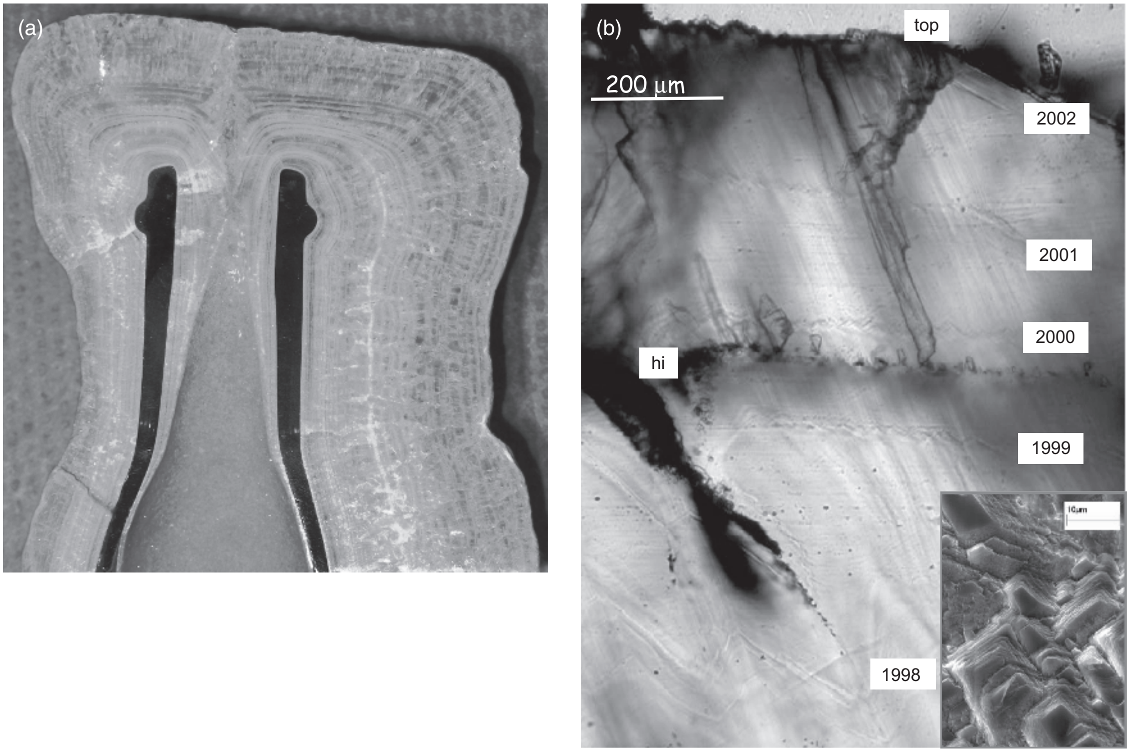

Fig. 1.5 (a) Sectioned speleothem that has grown around a bottle (black) in Proumeyssac Cave, Dordogne, France. Annual growth couplets in calcite represent alternations of clear calcite (darker) and calcite containing fluid inclusions (lighter). (b) Plate 7.2, thin section of the top of stalagmite Obi84, Obir Cave, Austria, showing growth from late 1998 to the end of 2002 by bright annual impulse laminae (plane polarized light). Black area lower left is air-filled inclusion and growth zones show that this occupies a depression on the growth surface. Hiatus surface (hi) represents a brief in-year pause in growth. Inset shows a scanning electron microscope view of the crystal growth surfaces (after Fairchild et al., 2010).

Because each radiometric data refers to one particular growth layer, what is then needed is an *age-model*: that is a continuous function of age versus distance on the sample. Fig. 1.6 illustrates data from several Chinese stalagmites. Stalagmite DA (Wang et al., 2005) displays nearly linear growth for the past 9000 years, whereas sb10 shows a slightly less regular growth with a hiatus of several thousand years. In both cases, the dates are close together and have sufficiently small errors that there is no ambiguity about the age-model, but in other cases there would a choice of how exactly to draw the lines between the dates (Scholz & Hoffmann, 2011). Sample D4 displays several step-changes in the rate of growth, but growth tends to be linear between these steps. All of these records are perfect in the sense that there are no *age-reversals* (stratigraphically younger samples with older ages). In practice, many samples are less ideal for dating because they can contain a lot of original detrital Th, as shown by the presence of ^{232}Th. Although this can be corrected, it is not always known which ^{230}Th/^{232}Th ratio to use for correction and so errors can be much larger and the data can display age-reversals. Other samples may be difficult to date simply because of low U content. Where annual laminae are present, the optimal strategy is to combine lamina-counting with the U-series chronology.

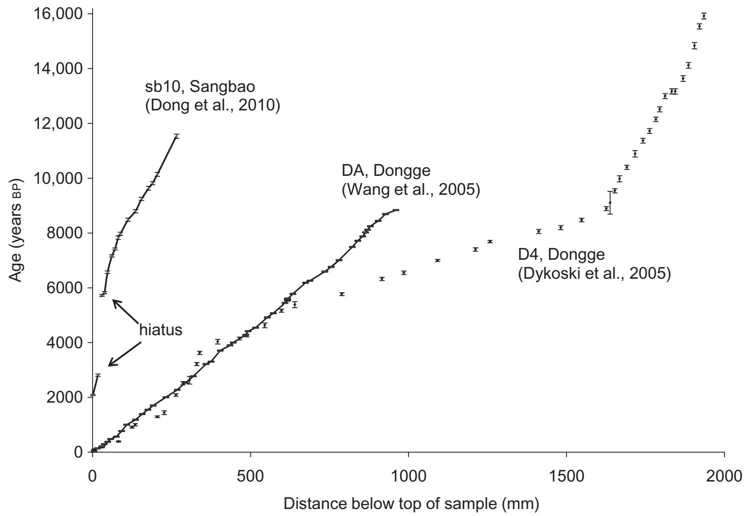

Fig. 1.6 Age models for some Chinese stalagmites included in the compilations of Plate 1.2 (Dykoski et al., 2005; Wang et al., 2005; Dong et al., 2010). DA shows continuous quasi-linear growth, sb10 illustrates a hiatus, and D4 illustrates data points defining a series of growth periods with differing linear growth rates.

1.1.5 What are the proxies for past environments and climates?

Proxies are parameters that can be measured in an *archive* (e.g. speleothems, ice cores, etc.) and which stand in, or substitute for, an environmental variable (e.g. mean annual temperature, seasonal monsoon intensity, or vegetation type). Examples of speleothem proxy parameters are growth rate, Mg concentration, $\delta^{18}O$ or $\delta^{13}C$ signature (see Box 5.3 for isotope definitions). In palaeoenvironmental analysis, the behaviour of proxy variables over time is used to interpret the changing environments or climates. In some cases this can be done quantitatively by means of an equation or process known as a *transfer function* (Fig. 1.7). Fairchild et al. (2006a) attempted to show systematically how the environmental signal becomes encoded in order to improve our understanding of how the proxy signals work. They drew attention to five realms in which signals were generated or modified:

1 Atmosphere (input of energy and matter; e.g. amount or $\delta^{18}O$ composition of rain, temperature variability).

2 Soil and upper epikarst (processes of organic decomposition, mineral dissolution, water flow and mixing; e.g characteristic $\delta^{13}C$ of vegetation, typical trace element to Ca ratio from bedrock dissolution).

3 Lower epikarst and cave (degassing and prior calcite precipitation, evaporation; e.g. changes in trace element to Ca ratios or $\delta^{13}C$ signature; influence of solution saturation state on growth rate).

4 $CaCO_3$ precipitation (partitioning or fractionation of elements and isotopes between water and $CaCO_3$ possibly dependent on growth rate or other effects; changes require use of transfer functions to derive water compositions).

5 Secondary change (e.g. change of aragonite to calcite; exchange of ^{18}O between calcite and water in inclusions).

In practice, certain parameters in particular contexts have proved especially valuable. The most commonly used proxy is $\delta^{18}O$ (McDermott, 2004; Lachniet, 2009), which has been found to be most powerful in cases where the changing atmospheric $\delta^{18}O$ composition of the original dominates the variation in $\delta^{18}O$ in $CaCO_3$; this is thought to be the case, for example, in the Chinese monsoon records in Plate 1.2. In other cases, the effects of temperature *per se* on rainfall composition and on the

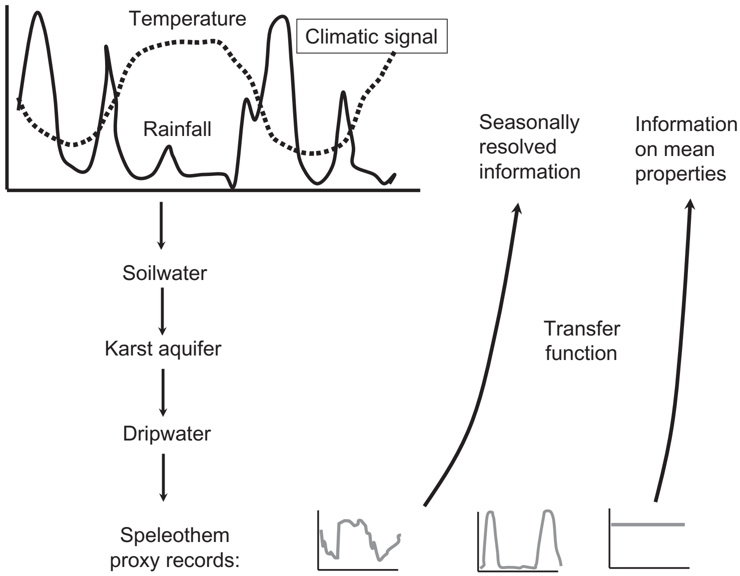

Fig. 1.7 Preservation of environmental (e.g. climatic) signals in speleothems and the use of transfer functions to inverse model the original environmental condition. An annual cycle of temperature and rainfall variability is used as an example. This is encoded in a speleothem by means of its proxy parameters. At one extreme, there may be a reasonably faithful representation, or a rectified version, of the original signal, whereas at the other, a characteristic constant value of a parameter is produced in the speleothem. Information on the original conditions can be recovered by reversing the coding process; where this is done quantitatively, an equation called a transfer function is used.

transfer function between water and speleothem composition can be opposed, and complicated by seasonal changes in rainfall composition: here difficulties in interpretation arise because of the local context. Likewise for $\delta^{13}C$, there are certain cases where the strong difference in isotopic composition between the aridity-tolerant C4 grasses and C3 vegetation is directly reflected in changing $\delta^{13}C$ signatures over time in speleothems. On the other hand, there are several other controls on $\delta^{13}C$, which has meant that such records are often left uninterpreted. Aridity can also be reflected in covarying $\delta^{13}C$, Mg and Sr signatures of speleothems, because of varying importance of prior calcite precipitation (Chapters 5 and 8). However, more generally many different patterns of variation of trace elements and isotopes with each other occur, only some of which are currently interpretable. Some specific signals such as high contents of colloid-transported elements make sense in terms of reflecting high rainfall and infiltration, but they are not expected to yield quantitative rainfall records. It is also worthwhile to mention that growth rate variations, determined from the thickness of annual layers, have been shown to reflect climatic variables in several cases and have been applied particularly prominently to modern climate calibrations and understanding the climatology of the last millennium (see Chapters 10 and 11).

1.1.6 How do speleothems compare with other archives?

The properties of Quaternary archives in general are well summarized in several texts (e.g. Bradley, 1999; Battarbee & Binney, 2008). The strengths of the speleothem archives lie in the following characteristics:

1 The common occurrence of continuous episodes of growth, thousands of years in duration, and the preservation of information representing timescales from days up to a million years. This is matched by ice cores, but other archives tend to have strengths either in their high resolution (e.g. tree rings) or long duration (e.g. deep marine sediments).

2 The excellent chronologies that can be obtained by U-series dating. This is far superior to the other proxies in which long records develop, although some other materials (e.g. coral terraces) can also be dated using the same techniques.

3 Speleothems contain several proxy parameters that can be used singly or in combination. Multiple proxies are being developed in numerous other archives too, including peat cores, lacustrine and marine sediments, ice cores, and indeed in individual components, such as foraminifera. It is the ability to record these parameters at very high-resolution which is a particular strength of speleothems and compares with marine coral records and tree rings for example.

4 They are widespread in inhabited continental areas and so their records have a direct relevance to regional climates and environments, and may also contribute to archaeological investigations. A limiting factor here may be the uniqueness of individual stalagmite samples, which makes conservation a concern (see Appendix 1), whereas other continental archives such as ice cores and lake records are typically more aerially extensive, as are flowstones.

5 Speleothems are physically and chemically robust and are relatively protected from erosion. This contrasts with other continental records, which are accordingly often not available before the Holocene. Certain lake records provide the strongest competition in this category and often have complementary information.

The weaknesses of speleothem science, in its current state of development, can be summarized as follows:

1 Insufficient inter-comparisons with other archives in well-constrained (e.g. co-located) contexts in order to improve the robustness of interpretation of both speleothem and other archives.

2 Insufficient understanding of the meaning of proxy variables. Many studies rely on one proxy and do not present data for others. Sometimes interpretations are unduly speculative or too simplistic in expecting a parameter to directly reflect a climate parameter.

1.1.7 What next for speleothem science?

In recent years there has been a rapid advance in analytical techniques to improve sensistivity and accuracy which has had the effect of reducing the destructiveness of sampling. Accordingly, even without further technical developments, there is an enormous amount to do in terms of applying state-of-the-art techniques to existing samples. In addition, there are many research frontiers, including the following:

1. The development of U–Pb methods to extend records throughout the Quaternary and into the Pliocene.

2. A more sophisticated understanding of how speleothem records contribute to understanding climatic drivers and teleconnections of climate.

3. Continued development of new types of proxy and the more effective use of multiproxy approaches.

We come back to these issues at the end of the book.

1.2 How is this book organized?

The first three chapters provide an introduction and context. Chapter 1 sets out our overall approach. The dynamic nature of speleothem-forming environments both on the human timescale and in deep (geological) time makes it attractive to apply the interdisciplinary concepts of system science as an aid to model development, as we discuss in the next section. We set out an explanatory framework for speleothems, firstly by introducing two new terms as metaphors: the *speleothem incubator*, the place in which speleothem growth is nurtured, and which lies within the *speleothem factory*, which also contains the delivery system for the raw materials for growth. We lay out the types of change occurring over different timescales and the parameters by which these changes are delivered; our task in this book is to provide a coherent connection between them. Chapter 2 focuses on the development and infilling of caves within human to geological timescales, emphasizing the various issues in carbonate geology. We outline the typical properties of aquifer carbonate rocks and their pores, the diversity of the cave systems, and the place of speleothems in the evolution and ultimate infilling or collapse of the cave. In Chapter 3, we show how the cascading and interlinked systems of climate, atmospheric moisture, soils and vegetation control many of the variables that determine the properties of speleothems. The different features of the climate system that might be recognizable in the past from speleothem studies are illustrated, as are the extent of variability of these climatic modes. We review how the biotic response to climate, modulated by the geological substrates, is reflected in the distinctive physicochemical properties of soils and their associated surface and subsurface ecosystems.

The second part of the book (Chapters 4–6) deal with the transfer processes in karst that characterize its dynamics and composition. Chapter 4 develops in detail the case for thinking of cave and karst processes as a speleothem life-support system: the speleothem incubator. We deal separately with the issues of water, air and heat movement in subsurface karst, illustrating the extent to which these can be understood quantitatively. We summarize by reviewing the integrative property of cave climatology and its influence on speleothem properties. Chapter 5 reviews the inorganic chemical components of the karst environment and its aqueous solutions in relation to speleothem composition. We deal with the fundamentally important carbonate system before considering each chemical variable, most of which have already been applied as palaeoenvironmental proxies when preserved in speleothems. Chapter 6 introduces the relatively neglected issues of the biogeochemistry of karst environments, illustrating the range of organic components that can be mobilized and incorporated into speleothems, and describing what we currently understand to be the important molecular transformations imparted on organic matter and the effects of organic assimilation, breakdown and transport on inorganic species.

The third part of the book focuses on the speleothems themselves, but we restrict attention to *calcareous* examples because these are overwhelmingly the types that are used in palaeostudies (Calaforra et al. (2008) discussed gypsum speleothems), using other analogues where appropriate. Chapter 7, on the architecture of speleothems, first discusses the fundamental controls on their shape and internal structure leading to a geometrical classification. A synthesis of their internal mineralogical, crystallographic and geometric structures sets an agenda for what needs to be explained in a given sample. Chapter 8 gives an overview of the geochemistry of speleothems, starting with the macroscopic setting, moving to the details of the water-speleothem interface and systematically describing the fractionations that occur between water and calcium carbonate during speleothem formation. This leads to two types of conceptual model: those that reflect patterns of seasonal change and those that reveal change in space within the cave environment. Chapter 9 on dating is particularly critical because a significant competitive advantage of speleothems in relation to other palaeoclimate proxies arises from one's ability to determine the absolute age of formation of a given speleothem layer. We deal systematically with the radiometric methods (those that depend on decay of specific radioactive isotopes), palaeomagnetism and annual-layer counting, before concluding with an analysis of age-depth models and uncertainties.

The most pressing argument for developing speleothem studies is their capability for contributing to palaeoscience (Roberts, 1998; Bradley, 1999; Batterbee & Binney, 2008), i.e. for recording information on past environments and climates, with an unparalleled combination of age-resolution and range of environmental proxies. Further, they are located in continental regions whose future climatic evolution is the subject of intense scrutiny. Study of the past boosts our confidence in forecasting (Skinner, 2008), particularly providing data for testing the *general circulation models* (GCMs) and other models used for prediction of future climate. We explore these aspects in the final section of this book, comprising Chapters 10–12. Chapter 10 tackles issues of calibration and validation of speleothem climatic proxies using young speleothems that have formed within the era of instrumental meteorological measurements. This is a rapidly developing field in which a range of approaches from related proxy materials are being used and there is a currently a major international effort to identify the key locations which can answer important climatological questions. In Chapter 11, we focus on the post-glacial Holocene Epoch, summarizing the range of relatively subtle climate drivers, key target time periods, quasi-periodic signals and model comparisons that have been investigated, with a particularly detailed focus on the last millennium. The Holocene is a period of immense archaeological interest as well as potentially one where disturbance by human activity is a key influence. Speleothems have already provided some intriguing pieces in the jigsaw, as well as more complete and meaningful records in some regions. Finally, in Chapter 12, we look at the longer Pleistocene Period (and beyond) in which the major Cenozoic glacial–interglacial cycles have occurred as has been documented primarily through study of ocean cores, ice cores, and some long lake and loess sections. Here speleothem records have been demonstrated in particular regions to be excellent integrators of aspects of hemispheric or global climate and have been used to refine timescales used by workers on other Quaternary proxies.

1.3 Concepts and approaches of system science

Systems science covers a diverse range of approaches to understanding the workings of natural and anthropogenic environments (Kump et al., 2009). Understanding the formation of speleothems requires a range of approaches, and systems thinking has two roles to play. Firstly, a qualitative systems analysis is useful to clarify our thinking about relationships of factors that may influence speleothem formation. Secondly, a quantitative approach allows us to test our understanding by building models that simulate aspects of the processes that lead to speleothem growth and initial steps in this direction are discussed in Chapter 10.

The science of systems has diverse origins, but threads derived from the mathematical modelling of regulatory processes (*cybernetics*, Wiener, 1948) and from organismal biology (von Bertalanffy, 1950, 1968) were highly influential in the post-war period. The detailed concepts of system theory were subsequently applied to geomorphic environments by the influential geographers Chorley and Kennedy (1971), contributing to geography's 'quantitative revolution', and were further developed by Bennett and Chorley (1978), although with no reference to karst.

Ford and Williams (1989, 2007) identified the development of structured networks of caves and drainage networks in unconfined karst as a case of a *cascading system*, which Huggett (1985) defined as '*interconnected pathways of transport of energy or matter, or both, together with such storages of energy and matter as may be required*'. Cascading systems are amenable to box-modelling (Box 1.1). In addition, we observe the development of morphological forms of defined geometry, which also show *self-organization* (the spontaneous development of ordered structures, such as calcareous deposits developing cascades of pools with rims). We might further aim to build a process–response model (Chorley & Kennedy, 1971) that brings together the quantitative mass or energy flow information with morphological parameters to give a more sophisticated understanding of the system.

A central issue in system science is the interactive nature of system components and in particular the occurrence of *feedback* behaviour, whether *positive* (reinforcing the applied change) or *negative* (resisting the change). Box 1.1 takes the well-known example of the carbon cycle to illustrate feedback concepts, showing how the issues can be downscaled to caves.

A variety of system concepts that refer to the nature of change over time (Chorley and Kennedy, 1971; White et al., 1992; Phillips, 2009) are illustrated in Box 1.1 and Fig. 1.8, and will be applied to karst in the next section. A system parameter may show an evolution over time, in which case it is described as *non-stationary* (Fig. 1.8a). Alternatively it may be *stationary*, that is, having a consistent mean value averaged over the period of interest.

Box 1.1 Box models and feedback

This box contains some basic material on flows of matter within simplified systems. Kump et al. (2009) give an accessible introduction, whereas Rodhe (1992) and Chameides and Perdue (1997) provide a more rigorous mathematical treatment.

A *reservoir* (also known as a box) contains matter (mass *M*) and can be considered homogeneous.

A *flux* (*F*) is the amount of material transferred from one reservoir to another per unit time.

Many systems are approximately *linear*, that is F_i (the flux or fluxes from reservoir *i*) is directly proportional to M_i, the mass in reservoir *i* (volume V_i can also be used instead of M_i for fluids at constant temperature and pressure).

If all relevant fluxes and masses remain constant over time, the system must be in kinetic or dynamic balance, and is said to be in *steady-state*. Under this condition, the underlying kinetic constants can be readily calculated.

A rate constant can be defined for each flux as follows:

$$F_{ij} = M_i \, k_{ij} \qquad \text{(I)}$$

where F_{ij} refers to the flux from reservoir *i* to reservoir *j* and k_{ij} is the corresponding rate constant.

$$1/k_{ij} = \tau_{ij} \qquad \text{(II)}$$

where τ_{ij} is the *residence time*, corresponding to the mean time that a given molecule resides in reservoir *i* before it is removed to reservoir *j*.

A relevant example is provided in Fig. 1.B1 by the comparison of carbon reservoirs and fluxes from the pre-industrial period (black), when the system was presumed to be in steady-state (black) with the changing situation in the 1990s (grey). Here the residence time (lifetime) of fossil fuels equals the reservoir size (3700 Gt C) divided by the current usage rate (6.4 Gt yr^{-1}), i.e. 580 years.

However, interpreting residence times when there are multiple fluxes out of a reservoir (i.e. multiple *sinks*) needs more careful discussion. For example, radioactive isotopes such as ^{14}C produced during atmospheric nuclear testing in the 1950s and early 1960s have provided important insights into element cycling. Half of the excess radiocarbon (^{14}C) disappeared in 10 years (Fig. 3.24). This observation has been used by climate-change deniers as evidence for the short lifetime of carbon in the atmosphere. Actually the *turnover time* of a reservoir (Rodhe, 1992), the time it would take to empty the atmosphere of ^{14}C if the sinks remained constant and the sources were zero, is even

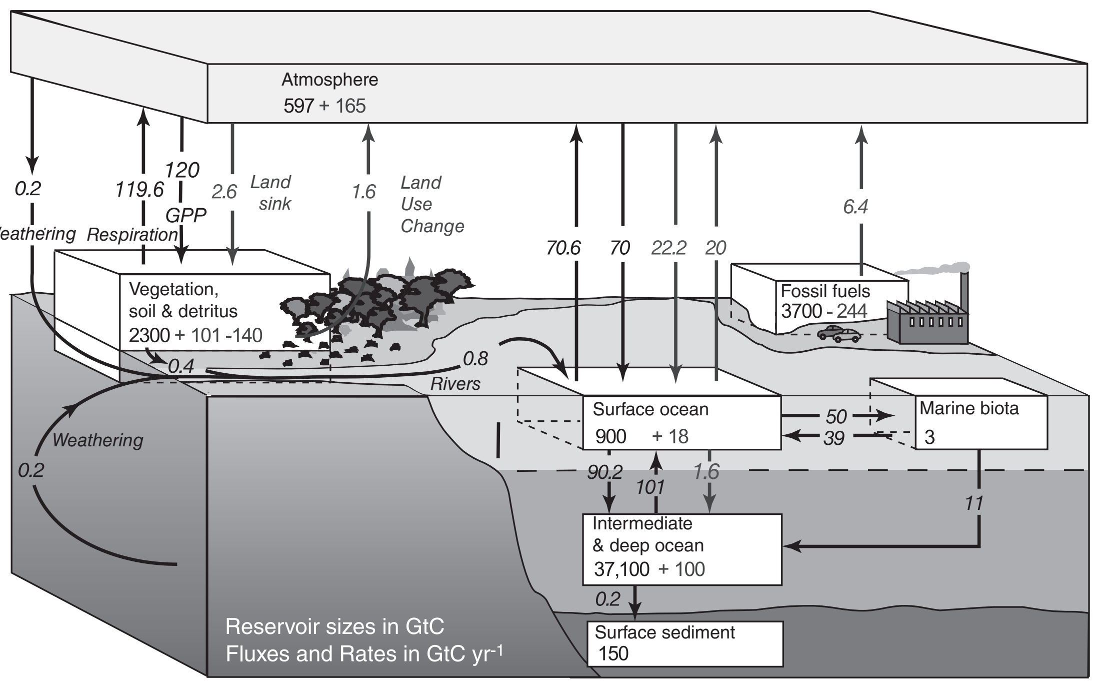

Fig. 1.B1 The carbon cycle (IPCC, 2007) showing reservoir (boxes) and fluxes (arrows). Figures in black are from the pre-industrial era whereas those in grey (red) are the changes resulting from human activity in the Anthropocene.

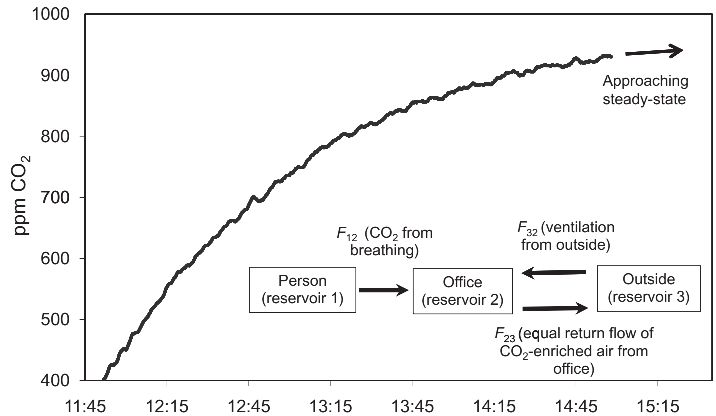

Fig. 1.B2 Results of an experiment to illustrate the concept of a system evolving to a new steady-state after a change in boundary conditions (see text for explanation).

shorter. In Fig. 1.B1, using the figures (in black) from the steady-state pre-industrial period:

Atmospheric carbon turnover time = mass/total outwards fluxes = 597/(0.2 + 120 + 70) = 3.1 yr.

The key point here is that ^{14}C is temporarily stored in vegetation and the surface ocean and is rapidly returned to the atmosphere, so that a more meaningful figure for the residence time of carbon is given by the ratio of the atmospheric mass (597 Gt C) to the 'permanent' sink in sediments (0.2 Gt C yr^{-1}), i.e. around 2500 years. Such a long period is the reason that our current carbon emissions are a cause for concern.

If the system is then perturbed by a change in the sources or sinks of matter, the rate constants can be used to estimate how the system will change. Using the pre-industrial figures and applying eqn. (I), the rate constant for transfer of carbon from the atmosphere to the surface ocean is 120/597 = 0.201 yr^{-1}. Given the enhanced mass of atmospheric C typical of the 1990s (597 + 165 = 762 Gt), application of this rate constant suggests that the flux of C to the surface ocean should then be 762 × 0.201 = 153 Gt C yr^{-1}. In fact the figure is just 92.2 (70 + 22.2), implying that more complex types of behaviour need to be considered, e.g. by more detailed consideration of the processes involved (Rodhe, 1992; see also the carbonate system in Chapter 5). Nevertheless, the approach clearly indicates the sense of change that is expected.

The above case example is from (large-scale) Earth system science, but there are analogous examples in terms of temporary sediment storage in geomorphic systems and in terms of carbon storage in soils for speleothem studies (see Chapter 8). Changing carbon dioxide concentrations in air can also be used to provide a good example of how to apply linear systems theory in a more local context. An experiment was conducted in which an investigator entered a previously well-ventilated empty office, closed the door and window, and monitored the changing carbon dioxide levels. The pattern of change is shown in Fig. 1.B2, increasing from an initial value close to that of the external atmosphere at around 400 ppm, but clearly flattening off after 3 hours towards a new steady-state value. This can be rationalized in terms of a box model (inset in Fig. 1.B2) with the person respiring CO_2 as reservoir 1, and the office (reservoir 2) exchanging air with the infinite external reservoir 3.

The mass of CO_2 in the office rises at the following rate:

$$dM_2/dt = F_{12} + F_{32} - F_{23} \tag{III}$$

It can be seen that the negative feedback that results in a stabilization of the carbon dioxide level is an inherent part of the structure of the system: as the concentration of CO_2 in the office rises, it is more efficiently removed in each parcel of air that exchanges with the exterior. This is a perfect linear system where the mass of carbon dioxide removed to the exterior is directly proportional to the mass in the room, because the rate of air exchange is constant. Ultimately a new steady state is reached where dM_2/dt is zero:

$$F_{12} = F_{23} - F_{32} \tag{IV}$$

F_{12} can be estimated from general adult physiology based on the typical CO_2 content (3.6%) and volume (8 litres) of exhaled breath per minute. With V_2 = 32,000 litres, this would lead to an initial increase in P_{CO_2} of 9 ppm min^{-1} whereas the actual observed figure was 7 ppm min^{-1} (the subject was a slow breather!). The rate of air exchange through the window frame was not known at the start of the experiment, but can be calculated as 420

(Continued)

litres min^{-1} by simulating the experiment stepwise on a spreadsheet (see www.speleothemscience.info) and adjusting parameters to fit the observed curve. Hence the approach allows a complete quantification.

Mass balance approaches have many applications in geomorphic and geochemical studies. One key example in cave environments is the CO_2 level in cave air because this directly control rates of speleothem growth, but normally varies seasonally. Figure 1.B3 illustrates the main fluxes of CO_2 in Ernesto, a small Italian cave (Frisia et al., 2011). The rate of decay of transient high CO_2 levels from visitors was used to estimate cave air residence times and exchange rates with the external atmosphere during steady-state summer conditions. The exchange times were used to demonstrate a large flux of carbon dioxide from air-filled fissures in the surrounding epikarst, much larger than that which could be released from dripwater. When considering the transition from summer to winter conditions, additional information is needed because all the fluxes are liable to change with changing temperatures and aquifer conditions. Although the system does not remain linear, the system structure remains unchanged, and enables quantitative solutions to be derived (Faimon et al., 2006; Kowalczk & Froelich, 2010; Frisia et al., 2011). A key feedback turns out to be a reduction in epikarst CO_2 flux because of an increase in water saturation of the aquifer in winter. These issues are explored further in Chapters 4, 5 and 8.

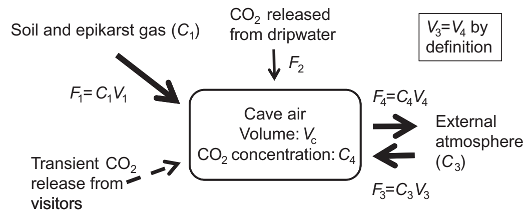

Fig. 1.B3 Structure of box model for carbon dioxide fluxes derived from a study of Ernesto Cave, Italy (Frisia et al., 2011). F, flux; V, volume; C, concentration.

Within this period, variation can be periodic (cyclic) as shown in Fig. 1.8b, or may be subject to non-periodic fluctuations. Cyclicity can be studied by a variety of statistical techniques that are introduced in Chapter 10. In Fig. 1.8b, the sequence is *deterministic* as the cycles are all exactly the same, whereas in nature there is a random component and hence the systems are non-deterministic or *stochastic.*

Where there is memory of previous system states, stochastic processes can give rise to apparent structured trends (Fig. 1.8c) without a corresponding systematic driving force. This is a really important concept for palaeoscientists because there is a strong temptation to over-interpret plots of variables against time (e.g. Fig. 1.8d), e.g. to consider all 'wiggles' on plots of past variables as meaningful (Fairchild et al., 2006a). The property is referred to as *autocorrelation* and statistical models can be created to describe it. Such models have an *order* which represents the number of previous states influencing the current state (Huggett, 1985). For example, in Fig. 1.8c, the upper line indicates a sequence of random numbers Y_1, Y_2, Y_3 . . . (within the range 0–1) whereas the lower (spiky) line is a random walk representing an example of an autocorrelative function (Z) of order 1 that is related to the random series Y at time t as follows:

$$Z_t = Z_{t-1} + Y_t - 0.5 \text{ (the term } -0.5 \text{ ensures that the series is stationary rather than rising).} \quad (1.1)$$

Figure 1.8c shows that apparent systematic variations are observed with the same degree of noise as the random number series. In real systems, it may be a major research task to distinguish purely random from systematic behaviour.

Stationary systems show a degree of resilience to a time-limited disturbance, as illustrated in Fig. 1.8d where the original system property is eventually recovered. An external forcing leads to system change after a *reaction time* and the system behaviour responds over a characteristic period known as the *relaxation time*. This contrasts with that shown in Fig. 1.8a where the system's response to a series of transient forcing events is to move closer to a *threshold* beyond which it rapidly moves to a different state.

In the geomorphic literature, the term *sensitivity* has been related to the probability that a persistent change will result from a disturbance (Brunsden and Thornes, 1979; Phillips, 2009); that is, a sensi-

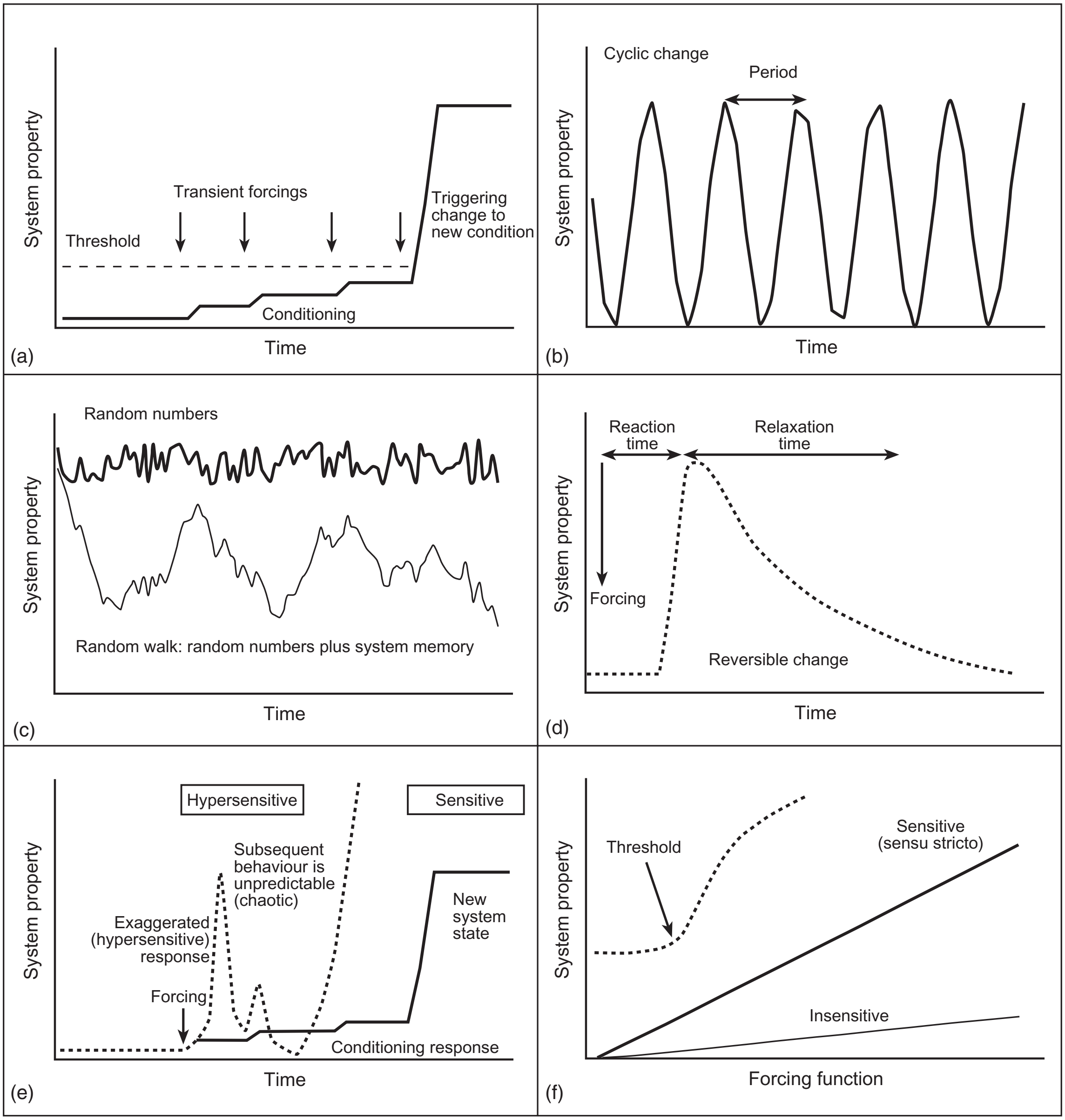

Fig. 1.8 Types of system behaviour. Diagrams (a)–(d) and (f) illustrate changes in a system property over time. In (a) there are a series of transient forcings; in (b) there is an implied cyclic forcing (although internal system dynamics can also give rise to this type of behaviour); in (c) the behaviour illustrates the addition of random forcings in a system with some memory effects; in (d) and (f) there is a single transient forcing with very different results. Diagram (e) illustrates the concepts of threshold and sensitivity in relation to a persistent forcing. See text for further explanation.

tive system (Fig. 1.8e) is opposite to the resilient system shown in Fig. 1.8d. In physical science literature more generally, sensitivity usually refers to a gradient between a system property and a long-lasting forcing function (Fig. 1.8f).

The behaviour of a system over time could show simple linearly scaled responses to disturbances (e.g. double rainfall produces double discharge), but real systems are more likely to be complex where exact physical solutions are unavailable: all we can then do is to ascribe probabilities for behaviour. A special case of complex systems are those described as *chaotic*, where there is hypersensitivity to the initial conditions such that the final result varies wildly and cannot be predicted (Fig. 1.8e). Phenomena driven by fluid turbulence, such as weather systems are of this type. Chaos theory and its relative catastrophe theory deal with nonlinear dynamics, that is, mathematically they refer to systems that cannot be described by systems of linear equations. Whereas a non-chaotic system can have a specific defined state (a defined solution to a mathematical equation) that can be described as an attractor, in chaos theory the (strange) attractors are fuzzy. For example, climate is the (strange) attractor for weather.

Catastrophe theory deals with system discontinuities, or *bifurcations*, where a small additional forcing can lead to a major change in system state that can be difficult to reverse such as the modelled behaviour of oceanic circulation in the Atlantic (Rahmstorf, 1995). The growth history of many speleothems reveals irreversible changes.

Currently, the dominant reference to systems in environmental science is Earth system science (ESS), which seeks to understand and predict the future changes in the interactions between the different major components of the Earth (geosphere, biosphere, atmosphere, hydrosphere and cryosphere), together with human interactions (Ehlers & Krafft, 2001). ESS makes most sense in the context of large-scale phenomena that ideally can be sensed remotely and which are sufficiently well understood as to be amenable to quantitative simulation: global biogeochemical cycling is an example (Chameides & Perdue, 1997). In this field, it is common to use industrial metaphors to describe the workings of the Earth, for example the text of Raiswell et al. (1980): *Environmental Chemistry: The Earth–Air–Water Factory*. The development of improved GCMs for climate underpins much large-scale research in ESS and is also highly relevant as a context for the interpretation of speleothem palaeoenvironmental research, as discussed in Chapters 10–12. To the extent that speleothems may contain encoded information on temperature, rainfall and vegetation, and are available over much of the land area of the Earth, they can also contribute to the testing of Earth system models, in which biogeochemical modules supplement the earlier purely physicochemical content of earlier models.

However, system concepts also apply at the local scale: what Richards and Clifford (2008) refer to playfully as LESS (local environmental systems science). Identifying the behaviour of individual geomorphic systems, their process-interactions, feedbacks and emergent properties provides a holistic basis for site-specific and generic studies of particular environments. As we will see, this turns out to be an enormously important insight for speleothem science.

1.4 The speleothem factory within the karst system

In this book, we propose a nested three-fold conceptual scheme (Fig. 1.9) in terms of the following:

1 The *karst system* as a whole where the emphasis is on its role as an aquifer and hence is on a much larger spatial scale than that relevant for speleothems.

2 A broader *speleothem factory* which operates as a cascading system, supplying the raw ingredients from the epikarst, soil-ecosystem, and adjoining atmosphere and transferring them to the cave environment. We develop this analogy further in this section.

3 A speleothem life-support system which we term the *speleothem incubator*. This consists of those parts of the karst and cave environment that maintain the growth of speleothems during a 'lifetime' when they are active. This maintenance includes the reg-

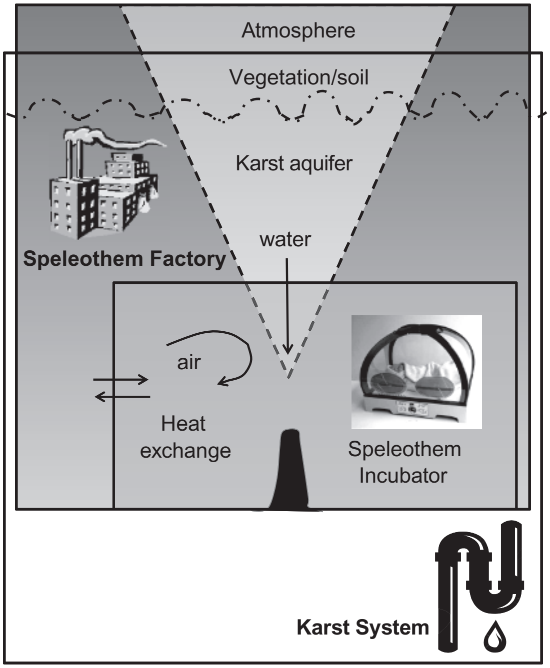

Fig. 1.9 The nested domains of (1) the karst system whose primary drive is to maximize the efficient transport of water; (2) the speleothem factory which supplies raw materials and generates speleothems; and (3) the speleothem incubator which is the cave and adjacent karst environment which maintain the conditions for speleothem growth.

ulation of air and water flows and is described in detail in Chapter 4.

Klimchouk & Ford (2000a) define the karst system as follows:

The karst system is an integrated mass-transfer system in soluble rocks with a permeability structure dominated by conduits dissolved from the rock and organized to facilitate the circulation of fluid.

White (1999) has lucidly summarized the key features of the system that need to be addressed in order to model water flow through karst aquifers. His outline structure of the system is illustrated in Fig. 1.10. Water is recharged at through autogenic sources (concentrated in closed depressions or as diffuse infiltration) or, from allogenic sources, at the surface (sinking streams) or sub-surface (perched aquifers). Karst is often rationalized as a triple-porosity aquifer (*conduits*, *fractures* and *matrix*) and, as shown by tracer tests, is typically organized into several groundwater basins in which discharge occurs in a single spring, or a small number of related springs. The underflow spring of Fig. 1.10 is the exit conduit at low flow, but overflow springs operate at high flow. Important nonlinearities are introduced by variable exchange between matrix and faster flow routes at different discharges, and by activation of overflow routes throughout the

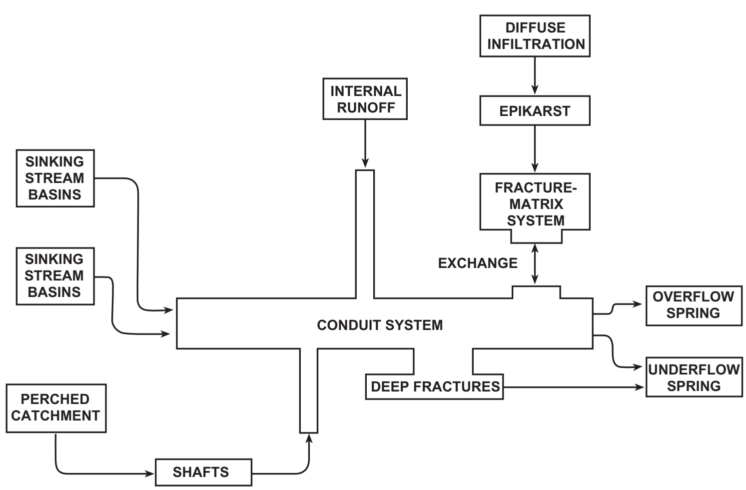

Fig. 1.10 Conceptual model for a karstic aquifer (White, 1999).

system. The key limitation is lack of knowledge about the three-dimensional layout of the drainage components, although constraints can be made by a range of observations including deconvolving discharge and chemical information at springs coupled with use of multiple tracers (see, for example, Perrin & Luetscher, 2008).

Hence it is clear that there exists a conceptual understanding of karst as an aquifer which marries with the key role of these aquifers in water supply. However, we do not currently have such an analysis for speleothem-forming environments. Here, we are often concerned with the life history of individual water droplets, while the major aggregated riverine flows do not leave a permanent chemical record inside the cave (although tufas may form externally; Andrews, 2006). Karst systems in the above sense are conceptually important for this book and so their development over time is discussed later (section 2.4). However, here our main focus concerns the conditions that determine the existence and properties of speleothems: this turns out to be something rather different.

The geomorphic literature stresses the role of physical forces in enacting change. Clearly landscapes are largely modified by the action of gravity in driving flows down slope and the dynamics of moving fluids in causing erosion and sediment transport. Aggradation occurs in areas where these forces are ameliorated. If we return to the industrial analogy, the construction of landscape involves a multitude of energetic grooving, planing and coating actions. This stands in distinct contrast to speleothem growth.

The growth of many speleothems, i.e. dripstones, involves the passage of pregnant waters drop-by-drop over surfaces accreting at rates that are so slow (micrometres to millimetres per year) that they barely overlap with those of industrial processes. Each speleothem body is unique, and in this sense they are analogous to archaeological artefacts. Indeed, they often co-exist with human artefacts with which they share great intrinsic value. This suggests that the manufacturing analogy would be that of a *workshop* where articles, tailored to individual customer requirements, are carefully constructed. Alternatively, the term *factory* is less ambiguous in terms of what work is conducted there, although there is a risk of implying mass production of uniform objects. In this book, we balance the unique aspects of individual speleothems against their generic characteristics, but in the issue of nomenclature, the generic has won out. Hence we refer to the speleothem-forming system, including the sub-system supplying necessary raw materials, as the *speleothem factory.* The great ceramic manufacturers of the Industrial Revolution were involved in a similar enterprise: the moulding from air, water and geological materials of arrays of objects, each readily identifiable by class, but with variety rendered by the detail of design or surface finish (Plate 1.1). Within the speleothem factory, a major sub-system is represented by the action of the karst aquifer and cave environment in regulating speleothem formation. This is the *speleothem incubator* of Fig. 1.9.

There are some conceptual similarities between the *speleothem factory* and the *weathering engine* of Brantley and White (2009). They are dealing with the weathering of silicate rocks in regolith and soil profiles in what has become known as the *critical zone.* Their *weathering engine* system encompasses the forcings, processes and consequences of this weathering. Carbonate rocks differ, however, in their more rapid growth and dissolution kinetics (Chapter 5) and the permeable nature of the epikarst compared with most regoliths.

In the following sections, we present an overview of the development of karst and speleothem-forming environments on different timescales as a summary account of the phenomena that are discussed in detail later in the book and hence it is expected that the reader may wish to return to this chapter and re-read it more critically once they have familiarized themselves with the evidence.

1.4.1 Long-term change

The speleothem factory represents both a distinct spatial zone within the karst system and a distinct stage or stages in the long-term evolution of cave environments (Figs. 1.11 and 1.12e). Speleothem formation is inhibited during the active stages of cave dissolution and occupancy by streams (Fig. 1.12a, b). A relative fall in base level arises natu-

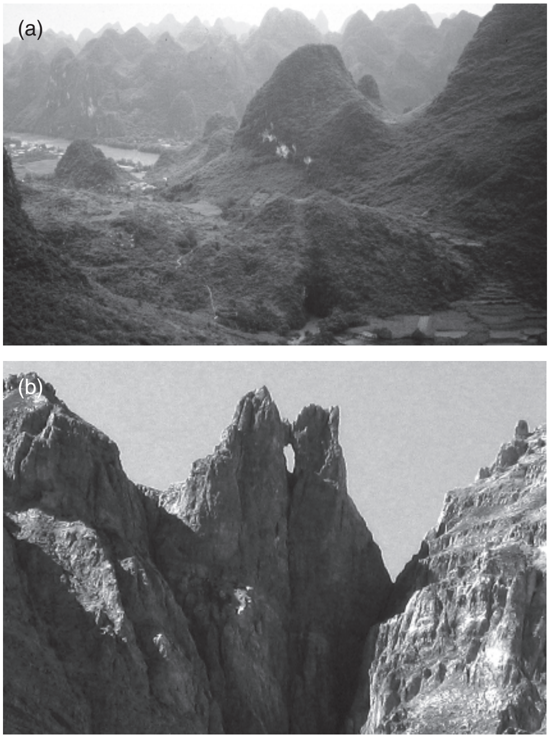

Fig. 1.11 Karst phenomena and high-relief topography. (a) Clon Cave, Guangxi province (Guilin karst region), China, a region with significant tectonic uplift enhancing topography. Tributary river entering a karstic massif at the end of a blind valley. The river continues to flow along a low gradient to reach the Lijiang river on the far side of the hill. (b) Val di Tovel, Trentino province, Italy. Old relic cave close to the skyline, developed along a sub-vertical fault zone, and near the end of its lifetime.

rally by the progressive gravitationally controlled downcutting of subterranean streams in order to provide a more efficient drainage (Figs. 1.11 and 1.12c). Base-level fall could also be accentuated given tectonic uplift or climatically controlled lowering of the water table. This sets the scene for speleothem growth. Once water is free to seep into air spaces of cave chambers and passages and degas excess carbon dioxide, and once active streamflow ceases, speleothem formation can begin. Ultimately the cave passage will be filled by a combination of speleothems and clastic sediment, unless first eroded. Complexity is introduced by the relationships of cave passages to specific geological features or the land surface (Fig. 1.12). This is a long-term cycle, and Ford (1980) noted that 'caves are the longest-lived (least time-limited) elements in [karstic] landscapes' (Fig. 1.11b). Individual caves may continue to accrete speleothems over time periods of the order of 10^4–10^7 years. Some infilled caves have survived over a billion years in association with palaeokarstic surfaces within sedimentary successions (Kerans & Donaldson, 1988).

In systems terms, the long-term evolution of karstic environments is one of continued evolution through successive thresholds on spatial scales of 10^{-2}–10^3 m with details that are contingent upon the exact bedrock configuration in a given location and the succession of climatic and regional geological histories. The main feedback process relates to the evolution of cavity systems. Initially they are systematically enlarged, but may become partly or wholly infilled by sediment, hence feeding back on the pathways of water flow further up the pressure gradient. There is also a chemical aspect to this feedback through the formation of speleothems. During the progressive enlargement of cave chambers and passages, there is an increasing probability that there will be significant air space within them. As air spaces become connected to the surface, exchange of gas with the surface increases. Hence the excess carbon dioxide degassed from solution can be removed which allows speleothems to form. Carbonate precipitation in seasonally air-filled cavities can block flow routes and hence feeds back on the hydrological system. The key point about this speleothem feedback is that it is a natural consequence of the creation of caves and hence inherent to the development of karst systems.

1.4.2 Annual-scale behaviour

Now let us focus on the key parameters of the speleothem factory and the timescales over which they vary. We start by considering its functioning over a year. How would an observer living in a cave at constant temperature be able to keep track of the seasons? Based on results from increased cave monitoring activities over the past 15 years, it is now clear that air composition, water quantity and

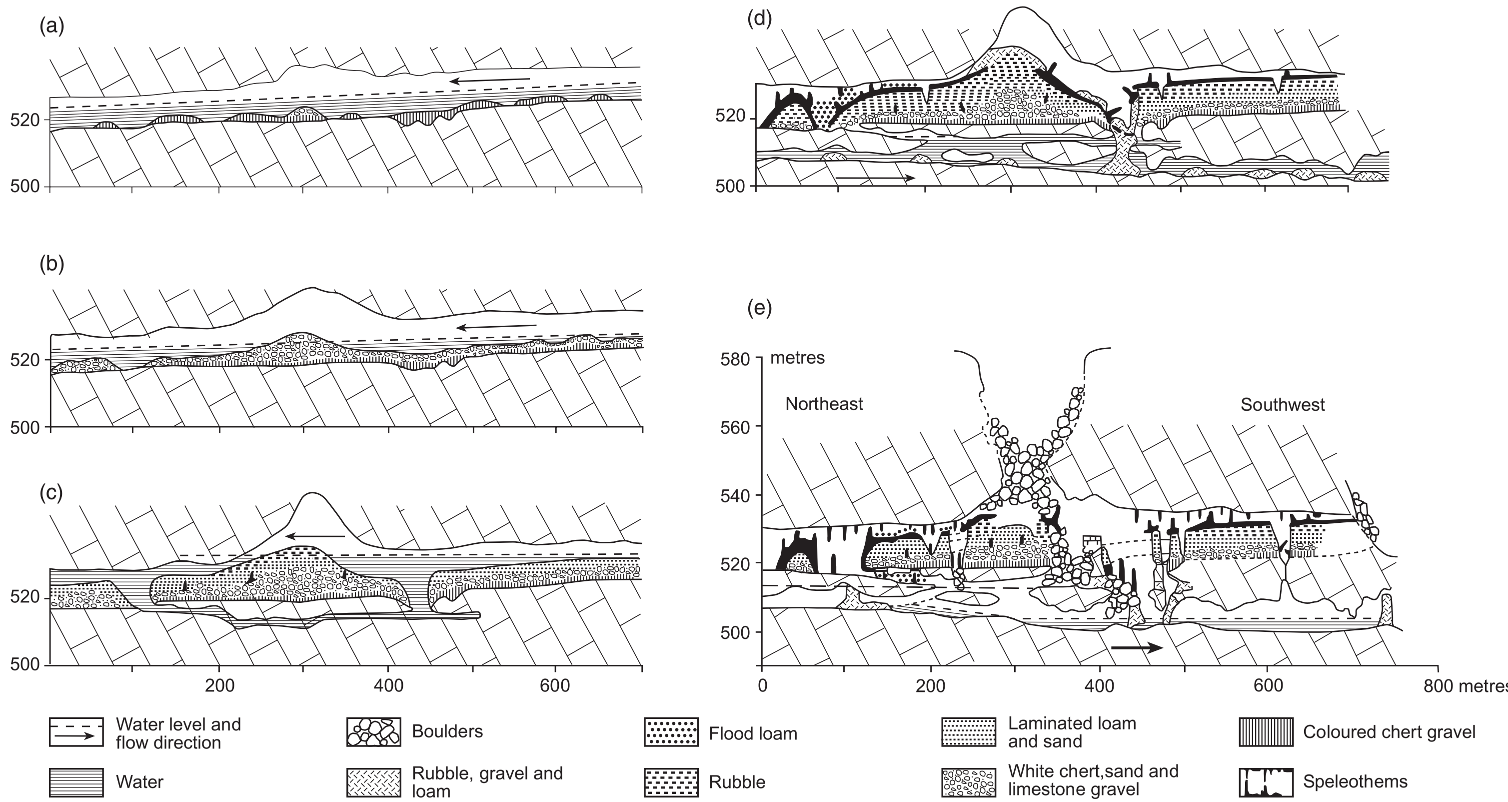

Fig. 1.12 Example of the reconstruction of long-term evolution of cave systems and their filling. (a)–(d) Development phases of the Otoska Jama cave, Postojna system, Slovenia, inferred from (e) which is the present-day configuration (after Gospodarič, 1976). A middle Quaternary sediment cone, resulting from roof breakdown, ponded water of the Pivka River and facilitated downcutting of the passage upstream of the cone. Downcutting was accelerated by continued sediment accumulation close to the ceiling of the upper passage. Phases of speleothem genesis are marked by flowstone layers and dripstone deposits with collapse of layers by erosive undercutting being evident. Reworking of sediment occurs, for example in association with debris movement from roof collapse and subsequent infill beneath the open aven (the Stora Apnenica doline).

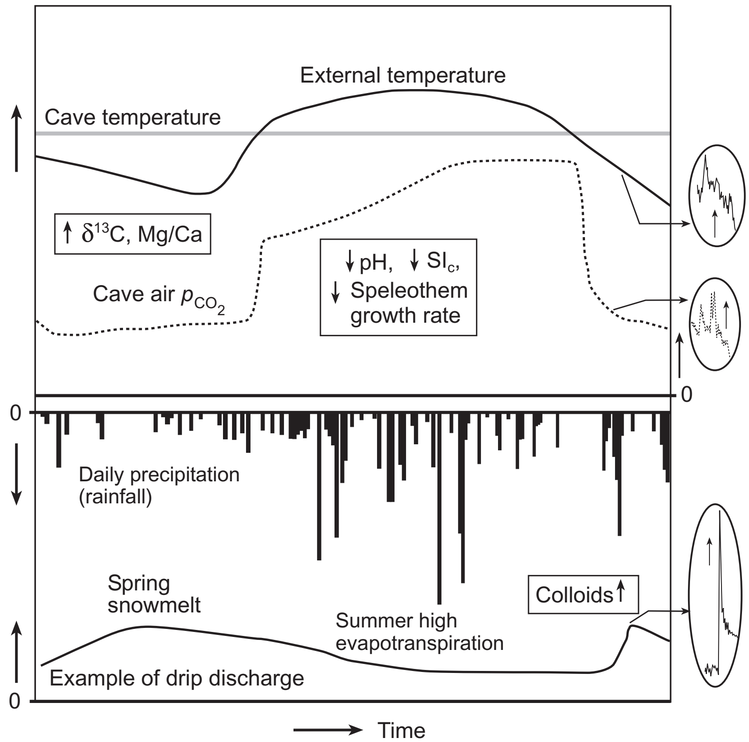

Fig. 1.13 Examples of some possible annual changes in parameters affecting the growth of speleothems, in this case based on study of Alpine environments including Miorandi et al. (2010). Insets (right) show detail of types of short-term (day to week-scale) behaviour. See text for discussion.

composition are all susceptible to annual changes: these are important to the operation of the speleothem incubator as is described in Chapter 4. Here we draw out some generalities (Fig. 1.13). The chemical terms given below are defined in Chapter 5.

It is increasingly recognized that the annual variation in temperature can have an over-riding control on the rate of air exchange between a cave and the exterior, although synoptic meteorology can also be important. Typically air-exchange is enhanced when the external temperature drops below the cave temperature and, unless there is an internal source of CO_2-rich air, the result is a lower P_{CO_2} of cave air in winter. Degassing of water in contact with the cave air is limited by its P_{CO_2}, and slow drips will attain a P_{CO_2} and pH in equilibrium with cave air. In summer, compared with winter, there is a lower supersaturation of the dripwater for calcium carbonate precipitation and hence slower growth rates. Greater degassing and $CaCO_3$ growth in winter is associated with an increase (both in dripwater and calcium carbonate), in $\delta^{13}C$, and also Mg/Ca further along a water flowline where $CaCO_3$ precipitation is occurring. Many speleothems display annual couplets in crystal texture that can be linked to bimodal cave conditions of this type.

Significant variations can also occur during the hydrological year. Each rainfall event is associated with a lag before the event water reaches a given position in the karstic aquifer and there will typically be an exponential decline in discharge over the relaxation time as in Fig. 1.8d. The cumulative impact of rainfall events (plus snowmelt during warmer spells in winter and spring) is smoothed by aquifer storage. A seasonal bias in infiltration should occur in all climate zones, although the gross year-on-year behaviour can vary to differing extents. The mean discharge of a drip is heavily dependent on the details of the water flowpaths, and is likely to relate more to the total filling of the aquifer than the time period of maximum infiltration. Even where this is the case, short-term increases in discharge can coincide with infiltration events. Where event water forms a significant proportion of dripwater, its seasonal variation can be reflected in that of the dripwater, e.g. in Alpine environments, winter rain and snowfall would be expected to be associated with light $\delta^{18}O$ signatures, and individual events could also be associated with reductions in saturation index and speleothem growth rate. Where saturation index is constant, increased water flux will result in increased speleothem growth rate. Perennially wet caves are characterized by a distinct flush of soil-derived colloids and trace elements in the autumn season. In contrast, caves associated with strong summer drought can be linked with enhanced degassing and calcite precipitation at an early stage along the water flowline during this season.

Dripwater hydrology can be approximated as a cascading system with linear characteristics, although nonlinear behaviour provides a more accurate description related, for example, to the complexities associated with air entrapment along flowpaths. The reaction times and sensitivity of individual drips are highly variable and thresholds for change are seasonally variable, being influenced by the overall state of water filling of the aquifer. Likewise, progress is being made in terms of quantifying the mass balance of air circulation as discussed in Chapter 4.

In summary, the strong drive of annual variations in temperature and quantity of infiltrating water, imprint themselves on the cave environment and dripwater characteristics, and hence on the associated speleothems. Different speleothems in the same chamber are likely to vary in terms of properties dependent on drip hydrology, but to display similarity in properties that are controlled by cave ventilation.

Inter-annual meteorological variation, both chaotic and associated with atmospheric modes of variability such as the El Niño/Southern Oscillation or the North Atlantic Oscillation, is inevitable. However, over a short timescale, the bounding system properties, as discussed in the next section, are stationary, or at least can recover their initial values; in other words, the changes are reversible. This may also apply to subtle changes over longer periods of time, perhaps up to hundreds of years, but in general the longer time periods usher in fundamental changes in system action as described in the next section.

1.4.3 Decadal- to multi-millennial-scale changes

Many of the parameters of the speleothem factory that can vary over these timescales are depicted diagrammatically in Plate 1.3 and listed in Table 1.1, together with their controlling variables and some relevant references. We have set the decadal scale as the lower limit for this section, because this is the minimum period over which we could recognize the onset of a long-term shift in climatic regime, such as that which occurred, for example, at the beginning of the Holocene 11,700 years ago (Rasmussen et al., 2006). However, many of the changes we discuss would require rather longer periods to develop. Individual speleothems are known to grow steadily for up to 10^4 years and more intermittently or with more strongly varying rates up to 10^5 years. At the 10^5–10^6 year timescale it is highly likely that structural changes to the karst system will occur, and in seismically active areas, flow rates and hence dripwater supply can change over periods more like 10^3 years. Long-term controlling variables of such time periods are listed in Table 1.1 for comparison with those that show changes over shorter periods.

The climatic controls are illustrated both directly and indirectly. A change in mean annual temperature will be reflected in a change in cave temperature (T) and a reduction in seasonal temperature contrast will tend to lessen the importance of cave air circulation (AC). A change in total atmospheric precipitation (P) will be moderated by any change in evapotranspiration (E), but $P - E$ changes will result in modifications to mean annual water infiltration to, and discharge from, the karst (Q). Increased rainfall intensity will result in relatively

Table 1.1 Parameters of the speleothem factory and their control over long and intermediate timescales. Sense of influence shown as (+) positive covariations, (o) complex, and (–) inverse variation.

	Parameter	Long-term controlling variables	Controlling variables for decadal–multi-millennial variation	Discussed in this book
AC	Strength of air circulation (with implications for contents of trace gases, aerosols and particulates)	Size of cave network (+), closeness to ground surface (+), surface relief (+)	Seasonal temperature variations (+); synoptic weather systems (+)	Sections 4.4 and 4.6
C	Soil CO_2 production	History of weathering and soil accumulation (+); position on catena (o)	Temperature (+), biomass (+)	Section 3.3
d	Thickness of water film on speleothem surface	—	*r* (–), *Q* (+), *x* (o)	Sections 7.2 and 7.3
dv	Carbon isotope composition of vegetation	—	Climate (C_4 vegetation in semi-arid grasslands) (o)	Sections 3.4
f	Drip fall height onto stalagmite or flowstone	Cave geometry	*h* (–)	Section 7.2
h	Height of speleothem	—	Duration and rate of speleothem growth (+)	Section 7.2
m	Vegetation biomass	—	Potential evapotranspiration (+)	Section 3.4
PCP	Prior calcite precipitation (calcite precipitation upflow of cave observation site)	Geometry of cavity systems in aquifer	Change in size of cavity systems (+); P – E (–)	Sections 5.5 and 8.5
P – E	Precipitation minus evaporation	—	Global climate (o); atmospheric modes (o)	Section 3.1
Q_1	Discharge via conduit flow	Existence of conduits	High-intensity rainfall events (+)	Sections 2.4 and 4.3
Q_2	Discharge via fracture flow (two alternative routes shown as 2a and 2b)	Aquifer properties	Relatively high-intensity rainfall events (+); seismic changes to aquifer properties (o)	Section 4.3
Q_3	Discharge via seepage flow	Aquifer properties	P – E (+), proportion of snowmelt (+)	Section 4.3
r	Radius of curvature of speleothem top	—	*Q* (–), *f* (–), *w* (–)	Section 7.2
R	Ratio in precipitation of stalactite versus stalagmite	Degree of degassing above cave	*PCP* (+)	Sections 7.2 and 7.3
RH	Relative humidity of cave air	Cave network geometry; location with respect to external atmosphere	*AC* (–), *P – E* (+)	Sections 4.2, 4.5 and 4.6
s	Soil depth and mineral properties	Long—term climate, bedrock properties, position on catena and aeolian input	Aeolian input (+), slope movement (o)	Section 3.3
S	Aeolian input	—	Aridity (+), glacial period (+), atmospheric circulation (o)	Section 5.3
T	Cave temperature	—	Mean annual external temperature (+)	Section 4.5
w	Width of speleothem	—	*Q* (+), *f* (+)	Sections 7.2 and 7.3
x	Crystal structure of speleothem	—	Supersaturation and impurity content of fluids (o)	Section 7.4

higher fluxes routed through conduit (Q_1) or fracture (Q_2) porosity, while increased snowmelt is likely to be reflected mainly in higher seepage flow (Q_3). The intensity and timing of rainfall may affect the entrainment of organic and other colloidal and colloidally bound components and their supply to speleothems. Meteorological/climatic changes tend to modify the mean and variation in the composition of $\delta^{18}O$ of atmospheric precipitation and hence of speleothems, but the atmospheric processes are complex and so will not be summarized here (see Chapter 3).

Many facets of the biota in the ecosystems overlying the karst are liable to change as the result of shifting climate. We have picked out the variables of total biomass (m) and $\delta^{13}C$ composition of biomass (dv) because these are parameters which have been of particular interest to palaeoenvironmental analysis. Distinctive biomarkers and relative abundances of different groups of organic entities and compounds will also be dependent on climate. The soil depth and composition (s), can vary, particularly with extraneous input from aeolian or upslope sources. The mean soil P_{CO_2} level (C) is an important parameter because it controls the maximum growth rate for speleothems; it is maximal at intermediate water saturations and seasonally higher temperatures.

Prior calcite precipitation (*PCP*) is a consequence of degassing of water into air spaces before the drip arrives in the cave under observation. Its sensitivity to dry conditions has been shown in several studies, but in detail depends on aquifer properties. The ratio of growth on stalactites and stalagmites (R) relates to *PCP*. A water that has already degassed will be ready to grow a stalactite (high R), whereas one that mostly degasses as it lands on a speleothem, will predominantly lead to stalagmite growth (low R). Factor *PCP* is associated with covarying increases in $\delta^{13}C$ and trace elements such as Mg in the stalagmite.

Various geometrical properties within the cave are interlinked. The width of the speleothem covaries with Q, but also tends to be higher because of splashing effects, when fall height (f) is greater. Increasing speleothem height (h) may significantly reduce f. The ratio h/w affects surface curvature (r) which, together with the surface roughness of the crystalline surface (x), affects the depth of the water film (d) on the stalagmite or flowstone surface, which in turn is a factor in speleothem growth rates.

The vigour of the air circulation affects the removal of internally generated gases such as carbon dioxide and radon (and its particulate daughters), but also the introduction of aerosols. More stagnant air can allow relative humidities to approach 100%, given sufficient inflow of water to the cave which leads to relatively slower growth, but closer to chemical and isotopic equilibrium.

In Plate 1.3, these parameters are shown to vary between the early stages of growth of a broad domal stalagmite in a humid climate (Plate 1.3a) and the later continued growth of a narrow form in a semi-arid climate (Plate 1.3b). In this case certain factors work together to reinforce a change, e.g. the reduction in water discharge (Q) and fall height (f) both promote the reduction in stalagmite width (w). Growth rate of the stalagmite is slower because of higher prior precipitation (P and R), lower discharge (Q) and more sluggish circulation (AC) as the result of higher mean temperatures (T) due in this hypothetical case to warmer winters.

So far in this section, we have discussed this qualitative systems analysis in terms of system parameters that in part are controlled by the setting of the cave in response to long-term geomorphological evolution and in part are linked to changing climates, as well as being inter-related. When studying speleothem samples however, the mental process needs to be inverted to examine the system from the point of view of the finished object in the speleothem factory. To what extent can we determine the processes and conditions of manufacture? Table 1.2 takes four properties commonly measured in speleothems (growth rate, $\delta^{18}O$ and $\delta^{13}C$ signatures, and content of the alkaline earth elements) and illustrates that for each of them, their baseline value is related to either geographic position or bedrock properties. However, each property is capable of being interpreted in terms of more than one controlling variable on both medium-term and annual timescales. Hence there is inherent ambiguity: this is like the situation where

Table 1.2 Identification of controlling variables for some parameters commonly measured in speleothems. For abbreviations, see Table 1.1. Sense of influence shown as (+) positive covariation, (o) complex, and (–) inverse variation.

Speleothem parameter	Long-term controlling variables	Controlling variables for decadal-multi-millennial variation	Annual-scale controlling variables	Discussed in this book
Growth rate	Karst aquifer properties	T, $P - E$	AC, Q	Chapter 7.2
$\delta^{18}O$	Location, altitude	$\delta^{18}O$ of atmospheric moisture (+); atmospheric dynamics (o), T (o)	Seasonal variation in weather systems (o)	Chapters 3.2, 3.3, 4.3, 8.4, 10.3, 11 and12.
$\delta^{13}C$	Location, altitude	m (–), dv (+)	AC (+), P (+)	Chapters 3.3, 5.4, 8.4, 11 and 12.
Mg, Sr	Bedrock composition	Aeolian input (o); P (+)	P (+), crystallographic factors (o)	Chapters 5.3, 5.5, 8, 11 and 12.

several manufacturers use different patented processes to produce a finished article with specified characteristics. Fortunately, this uncertainty can be resolved in many individual cases because of the dominance of one controlling factor, and in many cases there is a clear climatic implication.

In fact, over the past 15 years, remarkable evidence has been gained that speleothem factories on the decadal to multi-millennial timescale can behave in a coherent and reversible fashion. The strongest type of evidence is the presence of long-term cyclicity in measured parameters than can be matched with Milankovitch orbital variations, but sub-Milankovitch events are also often recognized, as is discussed in Chapter 12. On the other hand, many caves display speleothems which show extreme growth rate variations over much shorter periods of time, indicative of disruptions to aquifer properties, and curtailing their ability to record long-term processes. In between the very long-period Milankovitch variations and the expressions of the annual climatic cycle, are a range of stochastic behaviours for which we struggle, sometimes successfully, to separate random or karst-specific noise from an Earth system signal. In later chapters, we justify how these interpretations are made.

For future research, it is clear that work on monitoring cave environments, and associated soils and ecosystems, is vital to improve our understanding of speleothem-forming environments. There is also an important role for experimentation (e.g. Huang & Fairchild, 2001). In relation to the porcelain factory analogy, the experimentalist takes the goal beyond that of the antique expert studying the output of a past porcelain factory, towards that of setting up their own manufactory. Although timescales will limit what we can achieve with such control systems, they effectively complement the work of numerical models in the understanding of the workings of modern environments and the interpretation of proxy palaeoenvironmental records.