

Terra Incognita: Explanation and Reduction in Earth Science

Maarten G. Kleinhans, Chris J. J. Buskes and
Henk W. de Regt

The present paper presents a philosophical analysis of earth science, a discipline that has received relatively little attention from philosophers of science. We focus on the question of whether earth science can be reduced to allegedly more fundamental sciences, such as chemistry or physics. In order to answer this question, we investigate the aims and methods of earth science, the laws and theories used by earth scientists, and the nature of earth-scientific explanation. Our analysis leads to the tentative conclusion that there are emergent phenomena in earth science but that these may be reducible to physics. However, earth science does not have irreducible laws, and the theories of earth science are typically hypotheses about unobservable (past) events or generalised—but not universally valid—descriptions of contingent processes. Unlike more fundamental sciences, earth science is characterised by explanatory pluralism: earth scientists employ various forms of narrative explanations in combination with causal explanations. The main reason is that earth-scientific explanations are typically hampered by local underdetermination by the data to such an extent that complete causal explanations are impossible in practice, if not in principle.

1. Introduction

Earth science has received relatively little attention from philosophers of science. To be sure, plate tectonics and the dinosaur extinction event are frequently cited in the literature, but most of earth science is *Terra Incognita* to philosophers. Apparently, it is generally believed that earth science cannot offer much excitement from a philosophical point of view. One reason for this verdict may be found in the widespread idea that earth science is not an autonomous science but is easily reducible to allegedly more fundamental sciences such as physics and chemistry. While, in the past decades, serious

Maarten G. Kleinhans is at Universiteit Utrecht, The Netherlands; Chris J. J. Buskes is at Radboud Universiteit Nijmegen, The Netherlands; Henk W. de Regt is at Vrije Universiteit, The Netherlands

Correspondence to: Maarten G. Kleinhans, Department of Physical Geography, Universiteit Utrecht, PO Box 80115, 3508 TC Utrecht, The Netherlands. E-mail m.kleinhans@geo.uu.nl

doubts have been raised about the reductionist programme, and disciplines such as biology and psychology have thereby achieved the status of autonomous 'special sciences' (Fodor 1974; Dupré 1993), earth science has barely been mentioned in these debates, probably because the phenomena and processes in its domain are assumed to be purely physical and chemical and therefore easily reducible to lower levels.

The present paper makes a case for earth science as a subject deserving the attention of philosophers of science. Earth science has typical features that distinguish it from other natural sciences. We will highlight these features in the course of an investigation of the nature of earth-scientific explanations. The first goal of the paper is to answer the question of whether reductionism applies to earth science. Our answer to this question is tentatively negative (at least, not definitely positive). The second goal is to investigate the nature of explanations in earth science. We discuss whether narrative explanations are just descriptions, or sketches of causal explanations, or genuine alternatives to causal explanations. We find that earth scientists often confront underdetermination problems, which have serious consequences for the types of explanations that can be employed. The image of earth science that arises in the course of our investigation is one of a science with distinctive features that are philosophically interesting in their own right.

The outline of the paper is as follows. Section 2 briefly reviews the disciplinary aims of earth science. Section 3 addresses the reductionism issue. First, the nature of earth-scientific theories and laws is discussed. We find that earth-scientific generalisations fit the analyses of Beatty (1995) and Waters (1998), in the sense that they are contingent distributions rather than laws of a Nagelian type. Subsequently, we investigate whether emergent earth-scientific phenomena exist that are irreducible to phenomena at the chemical and physical level. Section 4 investigates whether earth science is characterised by special modes of explanation. We discuss whether narrative explanation is employed in earth science, and whether narrative explanation is reducible to causal explanation.¹ The notion of underdetermination plays a central role in our argument: we submit that earth-scientific theories and hypotheses usually are underdetermined by the available evidence, and that therefore complete causal explanations are out of reach. We conclude that earth science typically employs combinations of causal and narrative explanations.

2. Object and Aims of Earth Science

Earth scientists study the earth, that is to say, its structure, phenomena and processes, and their history. What exactly is it that earth scientists hope to achieve with their study of the earth? In his *System of the Earth* (1785), James Hutton, one of the founding fathers of earth science, proposed

to examine the appearances of the earth, in order to be informed of operations which have been transacted in time past. It is thus that, from principles of natural philosophy, we may arrive at some knowledge of order and system in the economy of this globe, and may form a rational opinion with regard to the course of nature, or to events which are in time to happen. (Hutton 1785, 2).

Apparently, Hutton saw the main aim of earth science as a historical description (acquiring knowledge of the ‘course of nature’), but the statement contains a suggestion that causal explanation and prediction are aims as well. Rachel Laudan, in her historiographical study of geology, distinguishes two aims:²

One is historical: geology should describe the development of the earth from its earliest beginnings to its present form. The other is causal: geology should lay out the causes operating to shape the earth and to produce its distinctive objects. (Laudan 1987, 2)

There are two related issues at stake here: first, the question of whether earth science is a historical or a nomological science, or a combination; second, the question of whether earth science aims for description, for explanation, or for both. The questions are related because historical sciences are typically associated with descriptive aims, while nomological sciences are associated with explanatory aims—the traditional distinction between historical description and nomological explanation. However, this is not a necessary connection: one might alternatively claim that historical sciences are explanatory, or that neither historical nor nomological sciences are explanatory. Incidentally, ‘real’ sciences may feature both historical description and nomological explanation, which implies that the distinction is one between ‘ideal types’.

McAllister (1997) formulates the first-mentioned contrast in terms of a distinction between ‘law-formulating sciences’ and ‘natural-historical sciences’. The former describe the world by means of laws of nature which abstract from nature (according to McAllister, laws of nature are ‘expressive symbols of the world’), while the latter produce literally true transcriptions of actual occurrences in nature. It seems that earth science is, at least according to Hutton and Laudan, both a natural-historical science and a nomological science. On the one hand, its goal is an accurate description of past courses of events, and on the other hand it is engaged in prediction (Hutton) and causal analysis (Laudan), activities which usually involve the formulation of laws.

The second question relates to a longstanding debate on the question of whether there exists a separate category of explanations, so-called narrative explanations, that can be provided by historical descriptions. Those who reject this idea either defend that historical sciences do not furnish explanations at all, or believe that their explanations are of the same kind as those in the physical sciences (choosing the latter option still leaves us with a choice between various rival theories of scientific explanation). In Section 4, we will inquire more deeply into the nature of earth-scientific explanation. We will argue that while earth science is partly a historical science, it does provide explanations because the historical descriptions provided by earth scientists are narrative explanations that are an integration of a common cause, robust-process explanations and actual-sequence explanations, observations, and background theories.

While earth science has, at least at first sight, a single object (the earth), a deeper look into any textbook reveals a bewildering variety of disciplines and approaches. Concepts and techniques are both developed within earth science and borrowed from other disciplines, such as logic, mathematics, physics, chemistry, biology, and computer modelling. A unified body of theory, topics, or techniques seems to be lacking. Many

practitioners of earth science viewed this as a shortcoming compared to the clear-cut identity and successes of sciences like physics and were led to reflect on the identity of earth science. Several authors suggest that earth science is divided into independent disciplines with different aims. For example, Rhoads and Thorne (1996) state:

There is a growing schism between those focusing upon reconstruction of the development of individual landscapes (often Quarternarists) and those seeking general principles governing landscape dynamics (most frequently numerical modellers, or process-oriented geomorphologists). ... the schism, if allowed to grow ..., could promote fragmentation of geomorphology. (Rhoads and Thorne 1996, x)

According to Rhoads and Thorne, process-oriented geomorphologists claim to be studying general physical processes, while quaternary geologists concentrate on the history of individual landscapes. Thus, they associate the former with causal explanation and the latter with historical description. Accordingly, the question of whether historical narratives can be reduced to general process-oriented explanations is a hot topic in institutes which home both disciplines (see Baker and Twidale 1991 for a vivid account of these debates). But in practice, the schism between geomorphology and quaternary geology and other disciplines is not that wide. Dorn (2002) gives a list of disciplines which play an important role within earth science: biology, civil engineering, geology, geophysics, sedimentology, human geography and physical geography (including geomorphology and biogeography), soil science, astronomy, chemistry, nuclear sciences, climatology, economy, and oceanography. All these disciplines appear in various combinations in the earth-scientific journals, suggesting that there is no apparent schism that divides disciplines into separate, independent sciences. The core of all environmental and earth-scientific disciplines is that they study various aspects of the (mostly but not exclusively inanimate) earth (and earth-like planets) in history and at present with a variety of concepts and techniques.

3. Autonomy of Earth Science

Can earth science be reduced to lower-level sciences like chemistry and physics? According to the traditional model of reduction (Nagel 1961), this requires that the laws and theories of earth science can be deduced from the laws of chemistry and physics, and this in turn requires bridge principles that connect the terms used in the different laws. There are two major obstacles to Nagelian reduction: (1) bridge principles often do not exist; (2) higher-level laws do not conform to the traditional criteria for lawfulness; in particular the requirement that laws are exceptionless generalisations (necessary for global reduction). Since many reductionistic scientists have the Nagelian reduction model in mind in their discussions of reduction, we will start our investigation into the autonomy of earth science by considering whether earth science is reducible in the traditional Nagelian sense. In Section 3.1, we discuss the nature of earth-scientific theories and laws, by investigating whether they conform to the traditional Nagelian conception of laws or to recent alternative conceptions, as proposed by Beatty, Mitchell, and Waters. Subsequently, in Section 3.2, we consider the status of allegedly emergent phenomena in earth science. We conclude, in Section 3.3, that earth

science is autonomous in the sense that it cannot be globally reduced in Nagelian fashion. In a more liberal sense, however, earth science can nonetheless be viewed as a reductionist enterprise.

3.1. Theories and Laws

What is the nature of the theories and laws that earth scientists actually use? First of all, it should be noted that because of the historical aim of earth science, its practitioners often use the term ‘theory’ in cases of hypothetical historical events. The most famous example is the ‘impact theory’ that is proposed to explain the K-T extinction: this ‘theory’ states that about 65 million years ago a meteorite collided with the earth, causing a radical climate change and a mass extinction of species among which were the dinosaurs. This hypothesis is a theory in the sense that it postulates an event or chain of events in the past (and therefore not directly observable anymore) that explains observed phenomena, but it is not a theory in the sense that it specifies laws or a general model for the explanation of phenomena.

Do earth scientists also use theories of the latter kind? An obvious candidate is the theory of plate tectonics, often hailed as the grand unifying theory of earth science playing a role comparable to that of natural selection in biology. Plate tectonics provides a mechanistic underpinning of continental drift (originally conceived by Wegener as a hypothesis explaining mountain building and the shape of the continents) and explains a host of other phenomena and processes, which cover almost all temporal and spatial scales relevant to earth scientists. Continental drift and mountain building are long-term processes, whereas earthquakes and volcanism are short-term phenomena which may readily occur within a human lifetime. For geologists and paleontologists, the long-term processes are relevant since they are concerned with the formation of deltaic deposits, the history of rock formation in a mountain range, the rates of tectonic uplift or downwarping in the sedimentary basin, the activity of faults in the area, etc. All these phenomena are covered by plate tectonics.³

As such, plate tectonics is the paradigm earth-scientific theory. But what is the precise nature and status of the generalisations that plate tectonics contains? Do they qualify as laws? The theory describes the formation and movement of plates, and postulates an underlying process in the earth’s inner parts (mantle convection) that is responsible for the forces that cause the plates to form and move. There have been attempts to model the mantle dynamics of other planets, which lead to the completely different pictures of rigid-lid mantle convection on Venus and mantle superplume dominance on Mars, agreeing with interpretations of various surface observations. Thus, plate tectonics provides a general model of crust formation and movement and of mantle convection that is valid for very different situations and planets. An example of a generalisation (a candidate law) within plate tectonics is: *Earthquakes are generated in the rigid plate as it is subducted into the mantle*. Is this a law? Not according to the traditional criteria, because it is not universal: it refers to a specific spatio-temporal situation, namely the earth as we know it today. It contains specific earth-scientific terms such as ‘earthquake’, ‘plate’, ‘subduction’, and ‘mantle’. If we translate these

terms into the language of chemistry and physics, we see that they refer either to contingent distributions of matter in Earth:

- mantle = the stony but slightly fluid layer surrounding the core of the planet, made of minerals rich in the elements iron, magnesium, silicon, and oxygen (as opposed to the core, which is the central metallic part of the planet);
- plate = broken piece of the rigid outermost layer of the Earth (at present, the Earth contains twelve plates).

or to processes related to the specific structure of the Earth:

- earthquake = failure of the plate when static friction is exceeded and a movement of one block with respect to the other block occurs, giving rise to oscillations or seismic waves;
- subduction = sinking of heavy material of the crust into fluid material, caused by the collision between two plates (which is in turn driven by convection in the mantle).

Although tectonics is applicable to other earth-like planets as well (see e.g. Van Thienen, Vlaar and van den Berg 2004), it is clear that its ‘laws’ differ from the laws of physics and chemistry in the sense that their validity is less universal and more tied to the specific constellation of the planet in question. For example, plate-tectonic models would not apply to the gas giants of the outer solar system. In other words, plate-tectonic generalisations such as that given above describe contingent regularities. At the earth-scientific level, the model does not contain laws in the traditional sense of universal, exceptionless regularities; it merely describes contingent phenomena, dependent on the particular configuration of the earth’s structure. These phenomena can be translated in physical and chemical vocabulary, as the ‘bridge principles’ above illustrate.

In this respect, plate tectonics fits the analysis that Beatty (1995) provides for the case of biological generalisations. According to Beatty (1995, 46), ‘all distinctively biological generalizations describe evolutionarily contingent states of nature’, and this implies that ‘there are no laws of biology’. Beatty specifies his ‘evolutionary contingency thesis’ as follows: ‘All generalizations about the living world: (a) are just mathematical, physical, or chemical generalizations (or deductive consequences of mathematical, physical, or chemical generalizations plus initial conditions), or (b) are distinctively biological, in which case they describe contingent outcomes of evolution.’ Beatty’s thesis has been challenged by various philosophers of biology, most notably by Sober (1997), Waters (1998), and Mitchell (2000). The thrust of their objections is that Beatty’s criteria for lawfulness are too strict, and that there is no reason to withhold the predicate ‘law’ from generalisations which are contingent to some degree and are less universally applicable than the laws of physics. In fact, as Mitchell (2000, 251–53) argues, even the laws of physics are ultimately contingent to some degree (but see Beatty 1995, 63, for an acknowledgement of this possibility). This criticism may be justified, but it has no immediate implications for the issue of reducibility of earth-scientific generalisations. As Sober (1997, S459) observes, Beatty’s thesis implies that “there is no additional layer of autonomous biological law that living things also obey.” But even if the objections

of Sober, Waters, and Mitchell hold water, and one would allow for biological laws, this does not entail that such laws are irreducible (cf. Waters 1998, 33).

We claim that Beatty's evolutionary contingency thesis applies equally well to earth-scientific generalisations such as plate tectonics. This entails that plate tectonics does not contain autonomous, irreducible laws. Moreover, even if a more liberal notion of 'law' is entertained, there is no reason not to assume that the plate-tectonic model and its laws can be reduced to lower-level physical or chemical theories. Earth-scientific generalisations, such as the cited example regarding earthquakes, describe contingent distributions and processes which can be reduced 'locally' because they can be exhaustively translated in physical and/or chemical terms.

But perhaps there are autonomous, irreducible earth-scientific laws to be found elsewhere. A candidate might be the set of principles used in the ordering of rock and sediment layers in historical explanations. The most important of these are the principles of superposition and of cross-cutting (Kitts 1966). Superposition means that a layer on top of another layer must have been formed *after* the lower layer was formed. The lower layer is a necessary temporal antecedent but not a cause of the formation of the upper layer. For instance, it is very unlikely that a sediment layer is deposited *below* another layer, because it would involve lifting (break-up, erosion, and new deposition) the upper layer, which obviously destroys the upper layer. Likewise, a layer cross-cutting other layers means that the cross-cutting layer was formed *after* the other layers. However, examples can be given in which these principles lead us astray. A set of layers may have been overturned in severe folding during mountain building, volcanic activity, or fluvial channels may form a cross-cutting layer at the same time of the deposition of the planar sediment, etc. Earth scientists are aware of these pitfalls and deliberately seek for evidence for such exceptions when applying the general principles. But the examples show that these geological principles cannot be regarded as laws in the traditional Nagelian sense: they have many exceptions. Consequently, these principles cannot be reduced globally, that is, according to the Nagelian model of reduction. However, this is an argument against Nagel's model rather than an argument against the reducibility of earth-scientific principles. As in the case of plate tectonics, there is a more liberal sense of 'local' reduction that applies to generalisations and principles that are less universally valid than physical laws.

We conclude that earth science does not have irreducible laws and that the theories of earth science are typically hypotheses about unobservable (past) events or generalised—but not universally valid—descriptions of contingent processes. In contrast to physics and chemistry but analogously to biology, an important part of theories of earth science consists in descriptions of contingent states of nature (Beatty 1995) or distributions (Waters 1998).

Incidentally, not only do earth-scientific generalisations resemble biological ones, but there is a close interaction between earth science and biology. Indeed, many earth-scientific phenomena would not exist without interaction with life. Figure 1a provides a random selection of examples of earth-scientific phenomena from very small to very large length and timescales. This typical reference to the length and timescales of such phenomena is discussed later. Figure 1b gives examples of earth-scientific phenomena

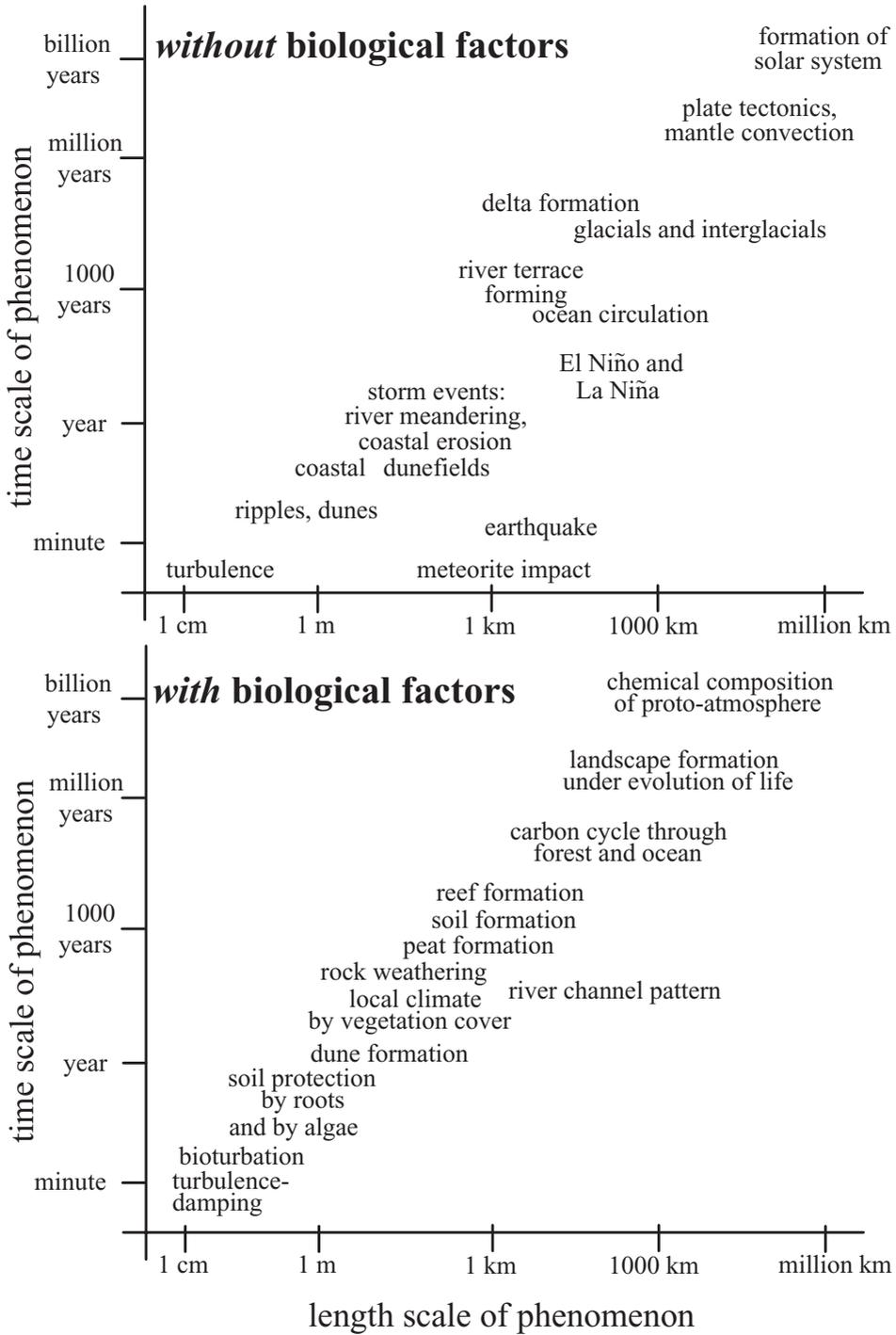


Figure 1 Examples of Earth-Scientific Phenomena without (Top) and with (Bottom) Interaction with Biological Phenomena.

which would not exist without biological elements. For example, the composition of the atmosphere of the Earth compared to that of Venus and Mars has much more oxygen and much less carbon dioxide. This so-called oxygen revolution was largely caused by photo-synthesising organisms (mostly algae) about 2 billion years ago. The presence of oxygen, in turn, led to increased oxidation of minerals, weathering of rocks, and chemical changes in the oceans. A second example is the variety of effects of life on rivers. Plant roots may stabilise the banks of rivers, which may actually cause rivers to change from a wide, shallow 'braided' planform with many mid-channel bars to a narrow, deep 'meandering' planform with one or a few channels only. Animals burrowing in or treading on the banks may initiate diversions of the flow and completely new courses of the river. At a larger scale, vegetation in the upstream catchment of a river may strongly damp the surface runoff from rain storms, leading to a much more regular water discharge regime than in unvegetated catchments, with all sorts of consequences for the morphology and geology.

3.2. Emergent Nature of Earth-Scientific Phenomena

In the previous section, it was argued that reductionism fails in the strong Nagelian sense, but a weaker (local) sense of reduction may still be applicable to earth science. Anti-reductionists might reply by invoking the alleged 'emergence' of earth-scientific phenomena as an argument against their reducibility. What precisely is emergence? According to Emmeche, Köppe, and Stjernfelt (1997, 83), it is "the idea that there are properties at a certain level of organization that cannot be predicted from properties found at lower levels". Wimsatt (1997, 373) states: "An emergent property is—roughly—a system property which is dependent upon the mode of organization of the system's parts". More specifically, Humphreys (1997, 341–42) lists the following possible criteria for emergence: emergent properties (1) are novel; (2) are qualitatively different from the properties from which they emerge; (3) could not be possessed at a lower level; (4) different laws apply to them; (5) result from an essential interaction between constituent properties; and (6) are holistic in sense of being properties of the entire system. Humphreys states that a property does not have to conform to all these criteria in order to be emergent. Wimsatt's definition is in line with Humphreys's criteria (5) and (6), while the definition from Emmeche et al. is consonant with criteria (1)–(4), especially (1) and (2). It thus appears that there is no precise, universally accepted definition of emergence. Moreover, there is not even agreement on the question of whether or not emergence is compatible with reduction: Emmeche et al. suggest incompatibility, but Wimsatt claims that emergence and reduction can go hand in hand. In the present section, we will investigate, using Humphreys's criteria, whether earth science provides cases of emergence and, if so, whether this provides an argument against the reducibility of earth science.

A first glance at earth science suggests that it is replete with emergence in one of the senses mentioned above: current ripples, rivers, deltas, volcanoes, mantle plumes, and continents consist of matter that is organised in such a way that novel, qualitatively different properties arise that do not directly follow from physico-chemical laws but

seem to comply with higher-level laws. Unlike biological phenomena, earth-scientific phenomena need not have a certain length or timescale, for many phenomena come in small and large sizes. For example, a tiny braided river may emerge in a few hours on a plot of one square metre. Actually, it can be produced in the laboratory under very simple, constant boundary conditions. This tiny river has many of its characteristics except its scale in common with braided rivers in the world, including the largest. The same is true for deltas, drainage networks, volcanoes, and many other small and large phenomena. The literature on emergent phenomena with scale-independent characteristics is extensive and focuses mainly on self-organisation, self-similarity, and chaos (see Stewart 1997, Smith 1998, and Ball 1999, for examples and details). For earth science, we identify two classes of phenomena.

The first class consists of self-similar, chaotic phenomena. These are characterised by a repeated basic pattern lacking a dominant length scale; the basic pattern occurs at a large range of length or timescales⁴ in the phenomenon, which is then called self-similar. The self-similarity at many scales is novel and qualitatively different from the microscopic properties of the constituent properties. Commonly, the largest scales occur much less frequently than the smaller scales, which is referred to as $1/f$ (f = frequency) scaling (Bak, Tang and Weisenfeld 1987). There are many examples of this class in earth science. Clouds and river drainage networks are well known for self-similarity (Ball 1999), as well as river-discharge records (Mandelbrot and Wallis 1969), sizes of avalanches on sand piles and of forest fires (Bak et al. 1987), elevation of landscapes and coastlines (Burrough 1981), iron-ore deposits, fault lengths, and fault surfaces (Lam and De Cola 1993). There is an interesting subclass: some phenomena are extremely sensitive to small perturbations, which means that they can evolve in completely different ways when initial conditions are even slightly changed. The best-known case is weather, but it is probably also true for climate, rock avalanches, earthquakes, and other phenomena. In some cases, self-similarity has been explained on the basis of lower-level physical laws (e.g. sand avalanches, crystal growth, and rock fault lengths), but in many other cases, this has not (yet?) been accomplished.

The second class consists of emergent phenomena with a dominant length or timescale. Such a macroscopic regularity emerges from microscopic physical or chemical processes. The microscopic processes may occur at a random or fractal length or timescales, but only one or two frequencies become dominant in the macroscopic pattern. The dominant frequency, or period or length, is in many cases enforced by a boundary condition that is independent of the microscopic processes. This type of phenomenon is emergent in the sense that it could be neither possessed at, nor predicted from, a lower level. A clear example is given by a tone coming from a violin. Moving the horse-tail hairs over the string provides a chaotic forcing on the string. The length of the string, however, leads to a dominant frequency of the tone. Other frequencies may be amplified as well and lead to the richness and other qualities of the sound, but there is one dominant frequency.

Such phenomena are common in earth science. For example, a turbulent water flow in a river provides a chaotic forcing to the sand bed below the current. Out of this chaotic forcing emerges a pattern: a train of large underwater dunes. Despite the

enormous variety in current velocity and sand characteristics, these dunes (when in equilibrium with the flow) have roughly constant length/height ratios and have a height that is about 20% of the water depth. This is an empirical fact for water depths ranging from 0.1 to 100 m and grain sizes from 0.4 to 100 mm. So, in practice, dune height and length are predicted from empirical relations with the water depth as the most important 'independent' variable. The dominance of a certain wavelength is sometimes used as evidence for identifying the process behind the self-organisation. In the case of the dunes, this is likely macroscopic turbulent pattern formation. Yet, although many attempts have been made to model turbulence over dunes, and to model incipient dune formation, it has not yet been proven possible to predict sizes and behaviour of real-world dunes from physically based models. Maybe it will be possible one day, but the simple empirical (statistical) predictors will be much simpler and faster to apply, albeit with uncertainties of about 50%. The dunes are therefore emergent in the sense of unpredictable dimensions from a lower level.

The way in which self-organising behaviour typically emerges in earth science is described succinctly by Werner:

First, development of collective behaviour by self-organisation reduces the very large number of degrees of freedom (for example, those characterising the positions and velocities of sand grains on a beach) to a much smaller number of independent dynamical variables (beach profile or shoreline position). Second, these variables evolve and interact nonlinearly to produce rich, emergent behaviour that is only weakly related to the original numerous degrees of freedom or the processes operating on them. (Werner 1999, 102)

Thus, the explanation of certain large-scale patterns in (in this case) coastal morphology cannot be found in the physical process of sand transport by waves and currents but has to be based on the 'independent variables' which interact nonlinearly. This does not imply that it is *only* self-organisation that explains these large-scale patterns; somehow, the sea bed is moulded into the pattern by sediment transport, which is a small-scale physical process. However, Werner argues that this physical process only determines the time needed to obtain the self-organised pattern; it does not determine the pattern itself.

For example, the self-organised criticality observed in braided rivers can be modelled with the nonlinear sediment transport process (where 'nonlinear' means that the transport rate increases with velocity or shear stress to a power larger than unity) and a number of macroscopic rules related to the water distribution over neighbouring grid cells (cellular automaton). A model of this system exhibits self-organisation in the sense that the flow is organised in a highly variable pattern with stable morphological (in this case, self similar) properties that are very similar to those observed in many natural rivers (Murray and Paola 1994).

3.3. Emergence and Reduction

It should be clear by now that many earth-scientific phenomena possess emergent properties. The question is now whether these properties are also irreducible. At

present, this question is hotly debated within earth science. Some earth scientists try to reproduce emergent patterns (e.g. a train of dunes or a braided river pattern) with simple models based on macroscopic rules only. By contrast, others attempt to reproduce these patterns with highly sophisticated mathematico-physical models, of which some elements and initial conditions have been simplified for reasons of computational efficiency. The former claim that the phenomena are irreducible, while the latter claim that given enough computing power and detailing of initial conditions, the phenomena can be explained (in a mathematical model) on the basis of physical laws only. Representing the former, Werner (1999, 104) defends a hierarchical modelling approach instead of a reductionist approach: “By isolating dynamics at different temporal scales, better predictability might be achieved with hierarchical modelling than in reductionist models.” Reductionists such as De Vriend (2004), however, claim that macroscopic self-organisation occurs due to microscopic processes and the laws of continuum flow mechanics. De Vriend and co-workers have reproduced emergent self-organised morphology in simulations of rivers, estuaries, and coastal seas by means of specific, physically based morphodynamic models in which the flow is modelled with continuum mechanics and the sediment transport with semi-empirical equations.

Both groups face problems. Burrough (1993, 102) has reviewed the use of self-similarity in landscape studies and concluded that although the estimated fractal dimension is a useful descriptive parameter, few surfaces seem to be truly self-similar; self-similarity occurs only over very limited ranges of scale. He explains this as follows: “Because natural phenomena are often the end result of a complex set of processes operating jointly over a wide range of scales, there is often no a priori reason why the resulting spatial or temporal variation should be self-similar” (Burrough 1993, 115). A fractal description gives no clue as to its real nature because there is a large number of processes operating and interacting simultaneously. Rather, the fractal distribution is used pragmatically as one of the available statistical distributions for representation of data. Even the ardent nonlinear dynamicist, Ball, warns against over-enthusiasm:

Some physicists have in recent years encouraged an unfortunate aspiration towards grand unified pictures [of pattern formation], but I fear we must accept that the world that we encounter ... is far too messy for that. ... [For instance,] the patterns of a river network and of a retinal nerve are both the same and utterly different. It is not enough to call them both fractal, or even to calculate a fractal dimension. To explain a river network fully, we must take into account the complicated realities of sediment transport, of changing meteorological conditions, of the specific vagaries of the underlying bedrock geology—things that have nothing to do with nerve cells. (Ball 1999, 252).

Thus, some aspects of river networks may be studied fruitfully with concepts such as self-organisation and fractals, but not all. Only few earth-scientific phenomena are truly fractal over at least a few orders of length or time which renders the fractal concept rather useless. Self-similarity without a dominant length-scale does not often provide clues about the underlying nature of the phenomenon, so it does not help to explain it. Self-similarity with a dominant length-scale provides clues on the forcing boundary conditions, but this is usually with large uncertainties, as in the case of the river dunes

or coastal sand bars where the boundary conditions vary strongly over time, but the phenomena do not.

Reductionist models of emergent phenomena, on the other hand, face serious computational problems and have problems with the comparison between model results and reality. Due to the necessary simplifications (like grids and discrete time steps), computational errors propagate and increase during long-term applications, leading to ‘exploding’ models. For example, in coastal morphological modelling, predictions for a few months to a couple of years ahead are demanded for coastal defence. But even with the most careful choice of parameters and specification of initial conditions, the currently available models eventually produce physically impossible results and crash (De Vriend 1997). The model schematisations and boundary conditions may be further simplified in order to prevent this, but then it is not clear anymore whether the modelled patterns that arise still resemble patterns in nature. This problem is not trivial; a cell-by-cell comparison of the emergent patterns between the model and nature is not useful (De Vriend 2004). In addition, it is not yet clear how to compare the more general pattern of model and reality quantitatively and objectively.

In this section, we have investigated the reducibility of earth science. First, we have compared it with the traditional account of reduction (Nagel’s model) and concluded that this account fails to apply because earth-scientific generalisations do not conform to the traditional criteria for lawfulness. This comes as no surprise, since these criteria are too strict to be able to deal with ‘laws’ and generalisations in complex sciences such as biology and earth science. We have discussed whether alternative accounts apply, and found that earth-scientific generalisations match Beatty’s evolutionary contingency thesis (Beatty 1995). This implies that reductionism is still a viable option (though not in the strict Nagelian sense), because Beatty’s account does not entail that higher-level laws are autonomous and is therefore compatible with reductionism.

Subsequently, we have investigated cases of emergence and shown that these are very common in earth science. But even this result does not entail that the question of their irreducibility is settled. Accepting emergence as a given fact, debates about reductionism continue among philosophers as well as among earth scientists themselves. Earth scientists disagree over the question of the irreducibility of earth-scientific phenomena. At present, some phenomena seem to be irreducible, while others have been explained with the laws of physics. In view of the recent reductionistic modelling, many more emergent phenomena may be explainable by the laws of physics in the near future. The fact that some earth-scientific phenomena are self-organised or fractal makes things both better and worse. In the case of phenomena with a dominant length or timescale, that dominant scale can be used as additional evidence for the process that causes the emergent pattern.

Meanwhile, philosophers have generally rejected the traditional Nagelian model of reduction, and rightly so, since it is clear that this model applies only to a very limited domain. In most higher-level sciences, including earth science, Nagel’s model is of little use. But this does not imply that in these sciences, reduction is completely absent.

Accordingly, we need an alternative account of reduction. Wimsatt (1997, S373) argues that “scientists in the complex sciences have compatible views of reduction and emergence”, and suggests that “a reductive explanation of a behaviour or a property of a system is one showing it to be mechanistically explicable in terms of the properties of and interactions among the parts of the system”. If one adopts this view of reduction, it appears that many emergent phenomena of earth science may be explained reductively in the future.

4. Explanation in Earth Science

In Section 2, we saw that earth scientists have at least two aims, namely to *describe* and to *explain* the (history of) inanimate processes on the earth. Indeed, it seems plausible that all sciences—including historical sciences such as geology—aim at explanation and understanding, over and above mere description, of the phenomena in their domain (see De Regt and Dieks 2005). But we have not yet answered the question of what the nature of earth-scientific explanations is: do they have a special status that distinguishes them from explanations in physics and chemistry? As argued in Section 3.1, earth-scientific theories are hypotheses about unobservable (past) events or contingent generalisations. How can such theories—which are largely of a descriptive or historical character—provide explanations? In the present section, we will try to answer this question by investigating the nature of earth-scientific explanations in detail. In Section 4.1, we argue that the historical narratives in earth science are in fact genuine explanations.

However, in almost all cases, these explanations are seriously troubled by problems of local underdetermination. In Section 4.2, we discuss how these problems play a dominant role in the practice of earth science. We suggest that earth science employs explanatory pluralism (although individual scientists of various disciplines in earth science may favour one type of explanation): there are several types of explanation, each with its own merits and its appropriate model. The combination of these types is often employed to overcome underdetermination problems (Section 4.3). In Section 4.4, we show that even the most reductionistic approach to causal explanation in earth science, mathematico-physical modelling, is strongly hampered by underdetermination problems, but can nevertheless contribute heuristically to narrative explanations. Finally, in Section 4.5, we show how two methodological strategies (Inference to the Best Explanation, Multiple Working Hypotheses) are employed in earth science to further overcome underdetermination problems.

We simplify the discussion of scientific explanation in a way that has no serious consequences for the case of explanation in earth science. We contrast ‘historical’ (or narrative) explanation with ‘causal explanation’, where we take the latter to be exemplified by the physical sciences. Causal explanation can take the form of a deductive-nomological explanation (à la Hempel) employing causal laws, or of a causal space–time description of events and processes (à la Salmon). For our purposes, differences between these two modes, as well as alternatives such as the unificationist view of explanation (Kitcher), can be ignored.

4.1. Are Historical Narratives Explanatory?

The issue of whether historical narratives are explanatory, and if so, how they are, is one on which opinions greatly diverge. Some authors claim that narratives are explanatory because they integrate an event into a bigger picture (Hull 1989). A narrative thus conveys some greater holistic insight, particularly in history and the social sciences. But others, following physicist Ernst Rutherford, sneeringly call historical sciences varieties of ‘stamp collecting’. Still others see a logical pattern in historical explanation that is fundamentally different from deductive-nomological and inductive-statistical explanation schemes. For human historical narratives, it has been argued that the plot, or argumentation structure of the story (e.g. perspective, ordering in time) conveys an explanation of the events in terms of their necessary (but not sufficient) conditions and their relations (see Von Wright 1971).

We will argue that there is more to the narratives than merely description. Earth-scientific descriptions are spatio-temporal narratives, and if there is anything explanatory to the narrative beyond the description, it must be because there is causal reasoning involved. Like evolutionary biology, earth science is partly historiography: it describes and tries to explain past events. The purpose of historical sciences is to provide a correct narrative of the sequence of past events and an account of the causal forces and antecedent conditions that led to that sequence (Lewontin 1994).

The historical-narrative approach provides two interesting kinds of explanations relevant for earth science. Sterelny (1996, 195) distinguishes *robust-process* explanations from *actual-sequence* explanations. According to Sterelny, both kinds of explanation are important, even for the same event, because they convey distinct breeds of modal information. Consider two explanations of the beginning of World War I. The actual-sequence explanation details the behaviour of Gavril Princip, who assassinated Franz Ferdinand. The robust-process explanation focuses on the political and armed alliances characterising Europe at the time. The actual-sequence explanation identifies the particular possible world that we inhabit: it tells us exactly how World War I was triggered. But it also leaves something out, namely that World War I was a war waiting to happen. If the war had not started with the assassination of Franz Ferdinand, it very likely would have started in another way. Thus, a robust-process explanation compares our world to others. The explanation identifies the feature characteristic of World War I worlds, and at the same time it reveals a host of counterfactuals. Robust-process explanations do not require full knowledge about all details and precise causal sequences of past events, yet they are as genuinely explanatory as actual-sequence explanations.

An earth-scientific example is the occurrence of a catastrophic landslide or mudflow, of which the precise timing is almost impossible to predict. We may model an observed flow in high detail with physical models. In retrospect, the initiation of the flow happened at that precise moment in time and at that location can be explained by, for example, heavy rainstorms. This actual-sequence explanation leaves out that the flow was waiting to happen because a certain amount of mud or rubble was on the verge of slope failure. If this specific heavy rainstorm had not initiated the flow, it would very

likely have been initiated by another rainstorm, or the spring melting of accumulated snow, or an earthquake. So, earth scientists not only model such flows with empirical or physical models but also map areas prone to mud flows ('waiting to happen') as a part of the robust-process explanation for these flows.

As far as descriptions and explanations of past events are concerned, Tucker notes that all historical sciences share two stages of development:

(1) A theory proves that similar evidence is not coincidental, but consists of similar effects that preserved information about their common cause or causes. (2) More evidence and theoretical background allow the probable reconstruction of intervening stages, informational causal links, between the common cause or causes and the similar effects. Further theories then allow the inference of properties of the common cause or causes from the information preserving properties of its effects. (Tucker 2004, 90–91)

An example is the theory of glaciation (ice ages). The common cause of glacials explains the occurrence of ice-pushed moraines, the absence of vegetation and many other observations, aided by theoretical background on the shear of ice-sheets on the soil, the effect of cold on vegetation and other theories.

In order to be genuinely explanatory, narratives in earth science need not be reduced to physical causal-law explanations. The practice of earth science demonstrates that a domain can be ontologically dependent on another deeper (i.e. physical) domain while at the same time being explanatorily autonomous. Hull (1989) has argued convincingly that in historical narratives, an event is not explained by subsuming it under a generalisation. Instead, it is explained by integrating it into an organised whole. In earth science, the ideas of continental drift and of glacials (ice ages) have played such a role. Mountain formation, earthquakes, volcanism, paleomagnetism, distribution of fossil remains, the form of the continents, and the varying ages of rocks are integrated in an organised whole, and all point to one and the same common cause: plate tectonics. Glacial landforms, erratic boulders, lack of vegetation in certain time periods, and evidence for a much lower sea level than today all point to the common cause of glacials. The specific physical processes of plate tectonics and glaciation are captured as contingent regularities with only local reductions to physics and chemistry (see Section 3.1).

4.2. Underdetermination Problems in Earth Science

While the historical narratives provided by earth scientists can thus be genuinely explanatory, they face another problem that we will discuss now, namely the fact that they are usually underdetermined by the available evidence. Philosophers distinguish between weak (practical) underdetermination and strong (logical) underdetermination. In the former case, there is insufficient available evidence to choose one theory over its rivals, while in the latter case, theory choice remains impossible no matter how much evidence is gathered. Many philosophers of science find weak underdetermination uninteresting, because it is 'merely' concerned with a practical problem that can in principle be resolved. However, we believe that weak underdetermination is actually

more interesting than strong underdetermination, precisely because it is a feature of scientific practice, especially in earth science.

Contrary to physicists, practising earth scientists very often face situations in which theories are underdetermined by the available evidence. In fact, it is hard to find papers that do not contain at least a paragraph on the way underdetermination was dealt with in practice (although it is usually not explicitly referred to as 'underdetermination'). Typical examples of underdetermination problems are detailed below (see also Turner 2005).

First, the timescale involved in shaping the earth is several orders of magnitude larger than the life of human observers or even written history. It is therefore problematic to detect and observe the long-term effects of slow processes that might be extrapolated to the past. Sometimes, controlled scale experiments are conducted in which detection and observation are less difficult than in nature, but this generates other problems that make a direct comparison with the real world more difficult.

Second, many earth-scientific processes and phenomena cannot (yet) be observed directly or even indirectly. Sometimes, a phenomenon eludes direct detection by instruments, for instance deep-mantle convection within the earth and other planets. Also, landforms and sediment deposits (with all the clues to processes and events in the past) often have been obliterated by erosion, mountain-building, or flooding. An additional practical problem is that current techniques often disturb the observed processes. For instance, the near-bed sediment transport on the sea bed and river bed is extremely difficult to measure without disturbing the local flow and sediment bed, while the disturbance affects the actual sediment transport that is being measured. So, in practice, considerable effort is focused on the techniques or measurement.

Third, many processes are intrinsically random or chaotic, and may be very sensitive to initial conditions. Often, earth scientists do not know which probability distributions or differential equations must be used to describe these specific processes. A veritable reconstruction of past events from the geologic record becomes extremely arduous because many events and phenomena are so complex that, in theory, an infinite number of possible laws and initial conditions could be involved. (This problem is called 'equifinality' in hydrology.) Noise and chaos lead to a certain uniqueness of geomorphological phenomena: duplications of events are seldom found, and the probability that the long-term river channel development in a particular river delta would take the same course when repeated is near-zero. This may be comparable to biology, where we also find many unique events, for instance the development of an individual organism, or the evolution of a particular species.

While complete causal (DN) explanations would be highly underdetermined, a narrative explanation does not require an exhaustively detailed set of observations and initial conditions. In many earth-scientific studies, initial and boundary conditions are implicitly (and without the details) given in the reference to the length and timescale of the phenomenon under study or the time period or study area, and in many other studies it is apparent from the context, such as the name of the scientific journal. For example, the Rhine-Meuse delta in the Netherlands has been studied by Berendsen and Stouthamer (2001) for the Holocene time period, which is the past 10500 years. The

point of this reference to time is to isolate the phenomenon and its causes from other phenomena by placing the latter in the description of the initial conditions. In this way, it becomes unnecessary to specify the whole chain of causes and events from the largest scales (long times past and large areas around the study area) down to the scales of the phenomenon itself. In the case of the Rhine-Meuse delta, it might be remarked that the delta is situated in the North Sea sedimentary basin, which is at a passive margin of the Eurasian continent. This description of initial conditions refers implicitly to the theory of plate tectonics with the terms ‘basin’, ‘passive margin’, and ‘continent’. It would be most impractical to include the theory of plate tectonics, other than in a descriptive way as initial conditions, in the theory of the formation of the Holocene Rhine-Meuse delta.

Common earth-scientific knowledge tells us that this approach is not only practical but also often defensible (Schumm and Lichty 1965). First of all, there is a correlation between the relevant length scale and the relevant timescale of most phenomena, so it is not necessary to study the evolution of a large object on a very short timescale or a small object on a very large timescale. For instance, current ripples, which have lengths of about 0.2 m, are formed and destroyed in minutes to hours, so studying ripples over decades is probably useless. Mountain ranges thousands of kilometres long, on the other hand, are formed in millions of years. In other words, it takes a much longer time to build or break down large things than to build or break down small things. The reason is that the range of energy available is between two close limits (compared to the extreme energy levels familiar to astronomers): a lower limit necessary to exceed thresholds such as friction or entraining boulders by stream flow, and an upper limit given by the maximum energy available on earth. Consequently, no significant change occurs to the whole mountain range in a period of hours, even though one plot on one slope may have changed significantly by a flood or by mass wasting. Therefore, it is commonly not useful to study large phenomena over very short periods or very small phenomena over large time periods. Notable exceptions are relatively unique events such as the impacts of large meteorites or earthquakes.⁵

4.3. Underdetermination and Explanation

How precisely do underdetermination problems lead to the impossibility of finding complete causal explanations and to the use of narrative explanations in earth science? Consider a typical earth-scientific research project that aims to construct a spatio-temporal description *and* an explanation of the course of the river Rhine in the past 10000 years (see Berendsen and Stouthamer 2001). Empirical evidence provides a host of indications that the river Rhine has flowed through the Netherlands and left its sediment in the past million years. In fact, the presence of bended features on the maps and air photographs, and the configurations of sand, clay, and peat layers in the ground, are all very well explained by rivers. The hypothesis that the Rhine has been present in the Netherlands is practically unassailable. But for earth scientists, this result is only the beginning, and not the end-point, of their investigations. What they really want is a description and an explanation of the course of events (their order in time and their

specific characteristics) that is generalisable to other, comparable phenomena (in this specific case, river deltas) in comparable circumstances.

For such an explanation, much more detailed evidence is needed to distinguish between competing theories of, e.g. river behaviour, mass extinction, plate tectonics, and climate change. Earth scientists are usually confronted with the predicament that they lack sufficient non-ambivalent data in order to construct a tenable explanation of the course of events. They have to infer from present situations to past ones, or from a limited set of observations to a hypothesis or theory. The empirical data gathered by earth scientists often leave room for a wide range of different, incompatible hypotheses. These hypotheses commonly are empirically but not logically equivalent; they cannot be true at the same time. In sum, their inferences are hampered by problems of underdetermination.

This conclusion is exactly opposite to that of one of the few philosophical papers on earth science (Cleland 2002). Cleland argues that historical sciences, among which she counts earth science, are not “epistemically inferior” to experimental sciences; they are based on a different but equally respectable methodology. According to Cleland, the predominant methodology of historical sciences is that of formulating various competing explanations (causes) of present phenomena (effects) and discriminating between them by searching for ‘smoking guns’. We agree that, as Chamberlin (1890) already pointed out, the methodology of ‘multiple working hypothesis’ is characteristic of earth science. However, Cleland invokes the alleged *overdetermination* of causes by their effects in order to conclude that the existence of ‘smoking guns’ is highly probable. By contrast, Cleland claims, experimental sciences manipulate present conditions (causes) in order to eliminate false positives and false negatives (effects), which result from the *underdetermination* of effects by causes.

There is a tension between, on the one hand, Cleland’s claim that ‘smoking guns’ can always be found and, on the other hand, her claim that various competing causal explanations of phenomena can always be formulated. It seems contradictory to assert that, on the basis of the presently available evidence, one can formulate various hypotheses and, at the same time, discriminate between them. We hold that Cleland’s claim that the evidence allows for more than one hypothesis is correct (this is in fact the thesis of underdetermination), but her argument of overdetermination of causes does not hit the mark.⁶ Consider a candidate causal explanation for the presently available evidence, e.g. regarding the courses of the river Rhine through time. This would take the form of a description of the initial conditions, say, 10 000 years ago, plus a set of causal laws that govern the dynamics of the system. The laws (L) are, as we argued above, derived from physics and chemistry (let us suppose they are given and deterministic); they are not the interesting part of the hypothesis. What really matters is the choice of the initial conditions (IC); these are actually referred to as ‘the cause’. But these conditions are never, and indeed never can be, specified *completely*: scientists are not Laplacian demons that have the ability to write down the state of the universe at a specific time in full detail—they merely give an incomplete description of a part of the universe (and this part is never a closed system). Scientists actually specify:

$$IC_{\text{specified}} = IC_{\text{complete}} - IC_{\text{unspecified}}.$$

Hence, while the ideal causal hypothesis (CH) would take the following form:

$$\begin{aligned} CH_{\text{ideal}} &= L_{\text{given}} + IC_{\text{complete}} \\ &= L_{\text{given}} + (IC_{\text{specified}} + IC_{\text{unspecified}}), \end{aligned}$$

the actual causal hypothesis is of the following form:

$$CH_{\text{actual}} = L_{\text{given}} + IC_{\text{specified}} = CH_{\text{ideal}} - IC_{\text{unspecified}}.$$

Consequently, CH_{ideal} would allow for an exact explanation of the course of events, and IC_{complete} would be overdetermined by its effects. But CH_{actual} does not allow for such an explanation, because of the residual $IC_{\text{unspecified}}$. The result is that the actually specified cause, $IC_{\text{specified}}$, is not overdetermined but underdetermined by its effects: there are many rival causes that are compatible with the available evidence (effects).

In the example of the Rhine river, the presently observed phenomena (e.g. configuration of layers in the ground) do not overdetermine the course of the Rhine in Holocene history. The deposits of older Rhine branches have been cut into shorter pieces by subsequent Rhine branches, and to connect the pieces together in the right way is a difficult puzzle requiring an enormous dataset of augers, radiometric datings, and other clues. The question of whether there was a river in this delta is heavily overdetermined by all the data, but this is trivial. The question of how rivers deposit sediment and how the auger data should be interpreted is also rather trivial. However, the explanation for the observed phenomena is not a trivial task.

Due to the underdetermination problems, earth scientists often cannot provide ideal, complete causal explanations. Similar epistemological problems occur in the best of scientific families. All sciences with 'historical' aspects face such predicaments, yet this does not hamper them to arrive at acceptable theories. As Tucker (2004, 27 ff.) has noticed, *consensus* is a key word here. Like any other science, earth science has a social dimension. When earth scientists reach a well-grounded consensus on a historical narrative, the best explanation is that they possess knowledge of the past. According to Tucker, a consensus which constitutes such knowledge must meet three criteria: the group that shares a consensus on beliefs should be uncoerced, heterogeneous, and sufficiently large. That is, the group should reach consensus on scientific grounds rather than through social pressure. Next, the heterogeneity of the group ensures that competing hypotheses are eliminated: uniformity of results in a heterogeneous group can confirm or refute hypotheses in controlled scientific experiments, or, as in earth science, by searching for smoking guns. Finally, the consensus should be sufficiently large to avoid accidental results because small groups can never be sufficiently heterogeneous to exclude hidden biases. Within earth science, these conditions are regularly met. Earth science, like any other discipline, has its outright disputes and different schools of thought, but on many important topics, researchers have reached a well-grounded

consensus. Hence, the explanatory narratives which earth scientists endorse are not mere fancies but constitute real knowledge of the past.

In practice, explanations in earth science are a combination of narratives and causal explanations. The narratives carry most of the explanatory power, for example of how the shifting of the Rhine branches over the delta plane depended on the sea-level rise and other factors. The causal explanation parts are, for example, applications of simplifying computer models to certain aspects of the delta development and are used as a test for parts of the robust-process descriptions. For example, the reaction of a branch of the river Rhine to the rising sea level can be modelled as a simplified system with physical laws. The outcome is a concave long profile along the bed of that branch. If the modelled profile matches the observed profile, then the rising sea level is a potentially good explanation for that part of the observations. The physical computer model is thus used to test whether the narrative does not conflict with the laws of physics. In the next section, we will consider the role of models in robust-process explanations, and the associated underdetermination problems, in more detail.

4.4. From Laws to Models: Further Problems of Underdetermination

Above, we assumed that the laws in causal explanations are unproblematic. But this is not the case in earth-scientific practice: the relatively simple laws of physics can seldom be applied directly to the initial conditions to check whether they explain the observations under scrutiny. Instead, physical laws are usually supplemented by intricate computer models. For instance, conservation of momentum and conservation of mass are simple physical laws that apply to fluid flow. These simple laws are the basic components of the Navier–Stokes equations that govern fluid flow. These equations are implemented in so-called ‘mathematical’ or ‘physical computer models’. For these models to work in practice, the equations have to be simplified, and discrete time steps and grid cells must be used to model the flow in space and time. The discretisation brings a host of necessary numerical techniques to ensure conservation laws and to minimise numerical (computer-intrinsic) error propagation. When initial or boundary conditions are specified for this model, certain laboratory or field conditions can be simulated. In this way, a model is used to test whether a hypothesis does not conflict with the laws of physics, and the model results can then be compared to the observations. But how can one be certain that a mismatch between model results and observations is not due to the simplifications and numerical techniques?

Both the initial conditions and the laws in the models are hampered by the epistemological problems mentioned earlier. Hence, unless the model is about a very simple aspect or field situation or experiment, there are always various sets of laws involved, of which some at best are derived from physics (but even then are highly simplified to allow numerical solutions). In these highly simplified cases, the various models may indeed converge to one solution, for instance in river morphology where various implementations of mass conservation and sediment transport models all give more or less the same results for sudden perturbations in the river bed. However, to solve Navier–Stokes for a *complete* river or coastal stretch without severe simplifications is

(as yet) impossible. Paradoxically, simplifications often make a comparison with field data more complex, because the data cannot be simplified in the same manner without adding uncertainties to the explanation. So, models are tested by, and calibrated on the datasets. But this does not solve the problems of underdetermination, for as Oreskes et al. explain:

Verification and validation of numerical models of natural systems is impossible. This is because natural systems are never closed and because model results are always non-unique. ... The primary value of models is heuristic. (Oreskes, Shrader-Frechette and Belitz 1994, 641)

Systems that are “never closed” exhibit the underdetermination of initial and boundary conditions, and the “non-uniqueness” refers to the equifinality problem. So, models cannot be used to hindcast or forecast the systems studied in earth science, unless they are calibrated on data from the past. But in the latter case, one can hardly speak of a true ‘physical’ model because the data carry a part of the explanatory weight.

And this is not the end of the problems for earth scientists. The example of the Navier–Stokes equations given above is a case in which it is well known which physical laws apply to the problem. But for many phenomena, it is not *prima facie* clear which physical laws apply. Moreover, the physical laws can be applied in many different ways: with many or few simplifications and with various numerical modelling techniques. In many cases, the laws are unknown, and only crude generalisations can be employed. In such cases, the explanatory power is increasingly carried by the data, and these ‘data models’ face well-known problems of limited predictive or extrapolative capacity. Because of this, verifications and validations of models are only of heuristic value. This is problematic. Suppose a number of competing hypotheses explaining a certain phenomenon have been conceived. An earth scientist could attempt to use the principle of parsimony (Ockham’s razor) to trim away physically and chemically impossible hypotheses. Since the laws of physics and chemistry come in the form of models, one cannot be certain that hypotheses are truly impossible or possible. So, in earth-scientific practice, Ockham’s razor is as subtle as a rusty breadknife.

Any earth scientist worth his salt knows this and exercises a healthy distrust of models. But there are nevertheless great benefits to be gained by modelling. First, the human mind cannot comprehend the results of complex sets of equations in space and time under certain initial conditions, but a model provides comprehensive results. Second, a scientist can manipulate a model in ways that are impossible in nature or even in the laboratory. Various scenarios can be studied and ‘what-if questions’ surveyed under the assumptions of the model, even if the modelled scenario is not what really happened according to the observations. Thus, these scenarios play an important role in robust-process explanations, both to extend the explanation and as counterfactuals in view of the observations. Third, a set of model runs may indicate (though not always prove) whether certain hypotheses are possible at all and whether they conflict with physical laws and mathematical constraints. The causal-law modelling in this sense is used as a test for the narrative robust-process explanations. Fourth, the comparison of various different models for the same phenomenon may indicate the robustness of the models: if the modelled phenomenon is similar in the different

models, then it is to some extent independent of the model schematisations. As Oreskes *et al.* rightly conclude, the value of models is primarily heuristic.

In line with the above, we believe that in earth science, both historical narratives and causal explanations are needed. Since the Earth's biosphere cannot be completely reconstructed in a laboratory, causal explanations have to be supplemented with historical narratives. The latter are needed to fill in the bigger picture. Hence, the two kinds of explanation are in fact complementary: causal explanations can be employed on a small spatio-temporal scale and may give further hints for narrative explanations on a much larger scale, such as the scenario-modelling approach outlined above, whereas such broadly construed narrative explanations may in turn provide more insight into specific causal processes. The border line between these two types of explanation is often unclear because they gradually merge into one another.

4.5. Inference to the Best Explanation

Until now, we have outlined the problems that earth scientists face. We do not want to suggest that these problems cannot be solved at all, however. In practice, earth scientists *do* come up with explanations, which they regard as more or less corroborated (cf. Section 4.3 on consensus). But how do the scientists arrive at these explanations? In this subsection, we will argue that most earth-scientific explanations are the result of abductive inference and that earth scientists typically rely on inference to the best explanation.

An important occupation of earth scientists is hypothesising about possible causes for the phenomena observed. Abduction in earth science is extended by a method already described by Chamberlin (1890) as 'the method of multiple working hypotheses'. A number of hypotheses are developed which potentially explain the observations. These can be processes that are known from observations, or 'outrageous' hypotheses of processes that supposedly occurred in past times but are no longer active or that seem to be in conflict with the laws of physics or chemistry. (Wegener's hypothesis of 'continental drift' is the most famous example of such an 'outrageous hypothesis': initially rejected as absurd, it is now accepted and supported by plate tectonics). The hypotheses are used to predict testable consequences, preferably for a wide range of different locations and with the use of different kinds of instruments. By contrasting and testing a number of (incompatible) hypotheses, a biased attempt at confirmation is prevented. Many different observations, derived from various techniques and from different locations, are used to test alternative hypotheses as rigorously as possible. It is helpful (but often not the case) if certain phenomena created in the past are comparable to phenomena created at present by known processes (Simpson 1975; Baker 1996). When these various data all point to the same (underlying) explanation or common cause, earth scientists accept this explanation as (tentatively) true.

For instance, the occurrence of about 23 glacial-interglacial couplets in the past 2.4 million years ('ice ages') is best explained by a combination of a unique setting of continents and astronomical influences on the global climate. There are cyclic variations in the inclination and direction of the earth's axis and the obliquity of the orbit,

for which the periods correspond with those of the glacial–interglacial cycles. These variations cause solar irradiation variations at higher latitudes, which leads to cooling. This, in turn, causes an increase in snow cover which leads to a larger albedo and hence causes more cooling. But this orbital forcing has been the case for much of earth’s late history while glaciation was not observed, so something more is needed. About 2.4 million years ago, Antarctica was detached from the South American continent. Circumpolar currents developed instead of currents circulating between the equator and the pole, which lead to thermic isolation and, given the polar position, cooling of Antarctica. The result was increasing ice coverage and hence a larger albedo, and consequently the whole planet cooled a few degrees. This cooling, added to the astronomically induced fluctuations, was enough to trigger the glaciation of higher latitudes. While this set of hypotheses is not without problems, an alternative set of hypotheses that explains these phenomena equally well is not easily conceived. For example, the fluctuation of solar activity due to its stellar dynamics (‘solar forcing’ on earth’s climate) is a competitive explanation for global temperature oscillations, for which effects have been demonstrated in the geological record. But there is no stellar theory that predicts the 23 rhythmic cycles, whereas the orbital forcing (systematic variations in the earth’s orbit) does. Most earth scientists believe that the former set of (triangulated) hypotheses is the best explanation for the observations, while the latter set of hypotheses is also true but only modify the effects of orbital forcing. This is an example of inference to the best explanation: the observed ice ages are surprising, but there are hypotheses explaining part of the evidence that would make the observations not surprising. So, the set of hypotheses is tentatively accepted as the explanation and further confronted with new evidence.

Inference to the best explanation has limitations (Lipton 1991). Foremost, abduction is a method to select *hypotheses* rather than at conclusive explanations. We probably will never know whether the best explanation is also the one and only true explanation. Perhaps the best explanation has not been discovered yet. This will sound familiar to earth scientists. They are often confronted with puzzles which ask for the most bizarre hypothetical solutions. In most cases, though, explanations are not truly discovered—they rather accumulate in time. Explanations in earth science are often so complex that they were not initially conceived as hypotheses at all. The uniqueness of events and phenomena in the geologic past asks for theories that are at first sight outrageous in view of our present experiences (Davis 1926; Baker 1996). The hypotheses that certain landforms are caused by ice-age glaciers, and that whole continents are drifting and colliding, were once outrageous. Over time, actual-sequence explanations were extended with further evidence, and robust-process explanations were formulated by extending the outrageous hypotheses and by adding causal explanation sketches (for example, by using mathematico-physical models). If the narratives survive tests against an increasing body of evidence, the hypotheses are more generally accepted by consensus as the best explanations, in accordance with Tucker’s account of theory development in historical sciences. As for the examples, the hypotheses of glaciation and plate tectonics are now generally accepted as the best explanations for many phenomena in geomorphology.

5. Conclusion

The aim of earth science is to provide descriptions and explanations and, if possible, predictions of (mostly but not exclusively inanimate) phenomena on the earth (and earth-like planets). In the present paper, we have highlighted some distinctive, philosophically interesting features of earth science, such as: the presence of emergent, but not necessarily irreducible phenomena, the important effects of underdetermination on theorising, and the explanatory pluralism employed by earth scientists. We have investigated the nature of explanation in earth science along two lines: first, by studying the prospects for reductionism by asking whether emergent laws or phenomena exist in earth science, and second, by inquiring into the nature of narrative (historical) explanation.

Earth-scientific theories do not contain *laws* in the traditional sense of universal, exceptionless regularities. In contrast to physics and chemistry but analogously to biology, an important part of theories of earth science consists in descriptions of contingent states of nature. Earth science is autonomous in the sense that it cannot be globally reduced in Nagelian fashion. However, we conclude that the traditional account of reduction fails to apply because earth-scientific generalisations do not conform to the traditional criteria for lawfulness; instead, they conform to more liberal views such as Beatty's and Waters's. If such a view is adopted, earth-scientific theories can be reduced locally because they can be exhaustively translated in physical and/or chemical terms. Accordingly, in a more liberal sense, earth science can be viewed as a reductionist enterprise.

Earth science provides many cases of *emergence*, but we have argued that this does not yet provide an argument against the reducibility of earth science. Earth scientists themselves disagree over the question of the irreducibility of earth-scientific phenomena. At present, some phenomena seem to be irreducible, while others have been explained on the basis of physical laws. Results of recent attempts at reductionistic modelling suggest that many more emergent phenomena may be explainable by the laws of physics in the near future.

While earth science is partly a historical science, it does provide *explanations* because the historical descriptions, or theories, provided by earth scientists are narrative explanations. Two kinds of historical-narrative explanations are applied in earth science: robust-process and actual-sequence explanations. In general, theories or narratives in earth science explain the geological phenomena by integration of a common cause, robust-process explanations and actual-sequence explanations, observations, and background theories.

Whether causal or narrative, earth-scientific explanations are strongly hampered by weak *underdetermination*, for which we have listed a number of typical examples. The empirical data gathered by earth scientists often leave room for a wide range of different, incompatible hypotheses. These hypotheses commonly are empirically but not logically equivalent; they cannot be true at the same time. Underdetermination entails the impossibility of constructing complete causal explanations, but narrative explanations remain possible because they do not require an exhaustively detailed set of observations and initial conditions. Rather, references to length- and timescales of

interest provide implicit limits for the amount of details needed and refer to relevant background theories. This reference also builds on the empirical fact that timescale correlates positively to length scale for most earth-scientific phenomena.

In practice, explanations in earth science combine narrative and causal explanations. The narratives carry most of the explanatory power in both actual-sequence and robust-process explanations. The causal explanation parts are, for example, tests whether the hypotheses do not conflict with physical laws. A conclusive test is, however, impossible due to the weak underdetermination problems. Two strategies are therefore followed: the method of *multiple working hypotheses* and *inference to the best explanation*. The uniqueness of many events and phenomena in the geologic past may require explanations that are at first sight outrageous in view of our present experiences. Over time, actual-sequence explanations are extended with further evidence, and robust-process explanations are formulated by extending the outrageous hypotheses and by adding causal explanation sketches (for example, by using mathematico-physical models). If the narratives survive tests against an increasing body of evidence, the hypotheses are more generally accepted by consensus as the best explanations. We believe that when earth scientists reach a well-grounded consensus on a historical narrative, the best explanation is that they possess knowledge of the past.

Acknowledgements

A paper on the philosophy of earth science cannot be written without discussions with many earth scientists, most of them in the Faculty of Geosciences at Utrecht University. They cannot be mentioned all. The discussions in the research group 'Knowledge, Normativity and Practice' of the Faculty of Philosophy, Vrije Universiteit Amsterdam, as well as discussions with Victor Baker, Gerbrand Komen and Arno Wouters at the first Dutch Philosophy of Geosciences Symposium (Utrecht, 2005) are gratefully acknowledged. Suggestions by James McAllister and two anonymous reviewers were very helpful in improving an earlier version of this paper.

Notes

- [1] We use the term 'causal explanation' generically, for the type of explanation employed in the physical sciences (in contradistinction to narrative, functional, and teleological explanation, for example). Of course, we are aware that this type of explanation can be further specified in a variety of ways, but for present purposes this may be ignored (cf. Section 4).
- [2] 'Geology' is meant by Laudan to include many disciplines like mineralogy, sedimentology, structural geology, geomorphology, geophysics, oceanography, paleontology, etc.
- [3] However, notwithstanding its great importance, the theory remains largely irrelevant for many earth scientists. For example, hydrologists studying local groundwater flow in deltaic deposits, and geomorphologists studying small-scale sediment transport processes on hill slopes, rivers, and coastal seas, will probably never refer to the theory of plate tectonics. For a climatologist, biogeographer, and oceanographer, the current layout of continents may be important as specifications in boundary conditions, but their interests do not extend to the explanation of these boundary conditions. Thus, although plate tectonics is important in earth science, it does not play the role of unifying theory comparable to natural selection in biology.

- [4] Note that it is ‘a large range of scales’ and not ‘all scales’. The self-similarity breaks down at both ends of the scales. The high end is the extent of the phenomena, which cannot be larger than the atmosphere in the case of clouds, or a continent in the case of rivers. The low end is the size of sand grains or molecules or atoms; the self-similarity cannot become smaller than this size. This need not be a problem, though. As Koperski (2001) argues, all representations (like laws) in science have this property. For example, fluid flow is governed by the (physical) Navier–Stokes equations. But this is continuum mechanics in which molecules or atoms, for which interactions would be described in discrete particle mechanics, have no place: “continuum models like those in fluid mechanics *ignore* the small-scale facts” (Koperski 2001, 697, emphasis in original). So, in self-similarity, we may also ignore the small-scale facts and the large-scale breakdown.
- [5] According to Tucker (1998, 65), unique events are topics of why-questions that radically underdetermine all potential explanations. He argues that there are no acceptable scientific theories that can explain unique events because such events occur once and only once. Yet, according to Tucker, we may question whether all alleged unique events are really unique. He therefore makes a distinction between *absolutely* unique events and *relatively* unique events. Relatively unique events are events that only *seem* really unique relative to a given context of scientific development and theoretical perspective. As science progresses towards more advanced levels of knowledge, the event may be discovered to be non-unique after all. For example, consider comets. These celestial events were once regarded as truly unique and individual, but Newton and Halley showed that many comets are recurrent and share the same physical properties. The mass-extinction of 65 Ma ago was caused by the impact of a comet on earth. Although once considered a unique event by earth scientists, evidence has been unearthed that a number of the mass-extinctions in the history of life on earth were caused by impacts of celestial bodies. It is an open empirical question whether absolutely unique events exist or not, but so far none has been identified by earth science.
- [6] Turner (2005) arrives at the same conclusion, namely that causal explanations are underdetermined by the data rather than overdetermined. His line of reasoning is different, however.

References

- Bak, P., C. Tang, and K. Wiesenfeld. 1987. Self-organised criticality: an explanation of $1/f$ noise. *Physical Review Letters* 59: 381–85.
- Baker, V. R. 1996. Hypotheses and geomorphological reasoning. In *The scientific nature of geomorphology*, edited by B. L. Rhoads and C. E. Thorn. Chichester, UK: Wiley.
- Baker, V. R., and C. R. Twidale. 1991. The reenchantment of geomorphology. *Geomorphology* 4: 73–100.
- Ball, P. 1999. *The self-made tapestry: Pattern formation in nature*. Oxford: Oxford University Press.
- Beatty, J. 1995. The evolutionary contingency thesis. In *Concepts, theories, and rationality in the biological sciences*, edited by G. Wolters and J. G. Lennox. Pittsburgh, PA: Pittsburgh University Press.
- Berendsen, H. J. A., and E. Stouthamer. 2001. *Palaeogeographic development of the Rhine-Meuse delta, the Netherlands*. Assen, The Netherlands: Van Gorcum.
- Burrough, P. A. 1981. Fractal dimensions of landscapes and other environmental data. *Nature* 294: 240–42.
- Burrough, P. A. 1993. Fractals and geostatistical methods in landscape studies. In *Fractals in geography*, edited by N. Lam and L. De Cola. Englewood Cliffs, NJ: Prentice-Hall.
- Chamberlin, T. C. 1890. The method of multiple working hypotheses. *Science* 15: 92–96.
- Cleland, C. E. 2002. Methodological and epistemic differences between historical science and experimental science. *Philosophy of Science* 69: 474–96.
- Davis, W. M. 1926. The value of outrageous geological hypotheses. *Science* 63: 463–68.
- De Regt, H. W., and D. Dieks. 2005. A contextual approach to scientific understanding. *Synthese* 144: 137–70.

- De Vriend, H. J. 1997. Prediction of aggregated-scale coastal evolution. In *Coastal dynamics 1997, conference proceedings*, edited by E. B. Thornton. Reston, VA: American Society of Civil Engineers.
- De Vriend, H. J. 2004. Prediction of morphological patterns. In *Marine Sandwave and River Dune conference II (MARID)*, Enschede, April 2004.
- Dorn, R. I. 2002. Analysis of geomorphology citations in the last quarter of the 20th century. *Earth Surface Processes and Landforms* 27: 667–72.
- Dupré, J. 1993. *The disorder of things: Metaphysical foundations of the disunity of science*. Cambridge, MA: Harvard University Press.
- Emmeche, C., S. Köppe, and F. Stjernfelt. 1997. Explaining emergence: towards an ontology of levels. *Journal for General Philosophy of Science* 28: 83–119.
- Fodor, J. A. 1974. Special sciences (or: the Disunity of Science as a Working Hypothesis). *Synthese* 28: 97–115.
- Hull, D. L. 1989. *The metaphysics of evolution*. New York: SUNY Press.
- Humphreys, P. 1997. Emergence, not supervenience. *Philosophy of Science* 64: S337–45.
- Hutton, J. 1785. The system of the earth, its duration, and stability. Reprinted in *Philosophy of geohistory: 1785–1970*, edited by C. C. Albritton. 1975. Stroudsburg, PA: Dowden, Hutchinson and Ross.
- Kitts, D. B. 1966. Geologic time. *Journal of Geology* 74: 127–46.
- Koperski, J. 2001. Has chaos been explained? *British Journal for the Philosophy of Science* 52: 683–700.
- Lam, N. S.-N., and L. De Cola. 1993. *Fractals in geography*. Englewood Cliffs, NJ: Prentice-Hall.
- Laudan, R. 1987. *From mineralogy to geology: the foundations of a science, 1650–1830*. Chicago: University of Chicago Press.
- Lewontin, R. C. 1994. Facts and the factitious in the natural sciences. In *Question of evidence: Proof, practice, and persuasion across the disciplines*, edited by J. Chandler, A. I. Davidson & H. D. Harootunian. Chicago: University of Chicago Press.
- Lipton, P. 1991. *Inference to the best explanation*. London: Routledge.
- Mandelbrot, B. B., and J. R. Wallis. 1969. Some long-run properties of geophysical records. *Water Resources Research* 5: 321–40.
- McAllister, J. W. 1997. Laws of nature, natural history, and the description of the world. *International Studies in the Philosophy of Science* 11: 245–258.
- Mitchell, S. 2000. Dimensions of scientific law. *Philosophy of Science* 67: 242–65.
- Murray, B., and C. Paola. 1994. A cellular model of braided streams. *Nature* 371: 54–57.
- Nagel, E. 1961. *The structure of science*. New York: Harcourt, Brace and World.
- Oreskes, N., K. Shrader-Frechette, and K. Belitz. 1994. Verification, validation and confirmation of numerical models in the earth sciences. *Science* 263: 641–42.
- Rhoads, B. L., and C. E. Thorne. 1996. *The scientific nature of geomorphology*. Chichester, UK: Wiley.
- Schumm, S. A., and R. W. Lichty. 1965. Time, space and causality in geomorphology. *American Journal of Science* 263: 110–19.
- Simpson, G. G. 1975. Uniformitarianism: An inquiry into principle, theory, and method in geohistory and biohistory. In *Philosophy of geohistory: 1785–1970*, edited by C. C. Stroudsburg, PA: Albritton, Dowden, Hutchinson and Ross.
- Smith, P. 1998. *Explaining chaos*. Cambridge: Cambridge University Press.
- Sober, E. 1997. Two outbreaks of lawlessness in recent philosophy of biology. *Philosophy of Science* 64: S458–67.
- Sterelny, K. 1996. Explanatory pluralism in evolutionary biology. *Biology and Philosophy* 11: 193–214.
- Stewart, I. 1997. *Does God play dice? The new mathematics of chaos*, 2nd ed. London: Penguin Books.
- Tucker, A. 1998. Unique events: The underdetermination of explanation. *Erkenntnis* 48: 59–80.
- Tucker, A. 2004. *Our knowledge of the past: A philosophy of historiography*. Cambridge: Cambridge University Press.
- Turner, D. 2005. Local underdetermination in historical science. *Philosophy of Science* 72: 209–230.

- Van Thienen, P., N. J. Vlaar, and A. P. van den Berg. 2004. Plate tectonics on terrestrial planets. *Physics of the Earth and Planetary Interiors* 142: 61–74.
- Von Wright, G. H. 1971. *Explanation and understanding*. Ithaca, NY: Cornell University Press.
- Waters, C. K. 1998. Causal regularities in the biological world of contingent distributions. *Biology and Philosophy* 13: 5–36.
- Werner, B. T. 1999. Complexity in natural landform patterns. *Science* 284: 102–104.
- Wimsatt, W. C. 1997. Aggregativity: reductive heuristics for finding emergence. *Philosophy of Science* 64: 372–84.