Philosophy of science: a practical tool for applied geologists in the minerals industry

J. Vann*[1,2,3,4] and M. Stewart[1]

For applied geologists working in the minerals industry the tasks of problem formulation, observation and data collection, interpretation and modelling invoke various philosophical considerations whether the practitioner is aware of them or not. A primary goal of applied geologists is to build models that accurately predict reality to an acceptable degree. In this paper, we describe the key philosophical frameworks proposed for conducting scientific investigations and relate them to the field of applied geology. We consider the very important differences in the types of problem confronted in experimental sciences (such as physics and chemistry) compared to the historical sciences, such as geology, where the processes studied are unique and only evidential traces of past events are available. The prediction quality of models is likely to be materially improved if the geologist is firmly and consciously practiced in the scientific method. In addition, if the predictions are framed and presented in terms of the underlying science, the quality of decisions made based on those predictions will likewise be improved. The implications for creating additional value to a project or operation can be very significant when geological models are constructed and used by a practitioner with an understanding of the philosophical basis of the activities constituting a scientific investigation. The method of multiple working hypotheses is particularly important when working in historical sciences. We argue that working within the framework of multiple working hypotheses can provide a valuable insurance against the adoption of, or persistence with, flawed models.

Keywords: Applied geology, Philosophy of science, Scientific method, Risk, Multiple working hypotheses, Falsification

'Remember that all models are wrong; the practical question is how wrong do they have to be to not be useful?'

Box and Draper (1987, p. 74).

'The demand that theories be highly falsifiable has the attractive consequence that theories should be clearly stated and precise. If a theory is so vaguely stated that it is not exactly clear what it is claiming, then when tested by observation or experiment it can always be interpreted so as to be consistent with the results of the test.'

Chalmers (1999, p. 67)

Introduction

This paper is written from the perspective of two experienced minerals industry geologists with an audience of industry peers and young professional geologists in mind. It is not intended to be an academic contribution to philosophy; however, the authors are convinced that there is a role for papers in technical journals to stimulate thinking about the activities we perform as scientists. There are very practical consequences for our discipline of being clear (or unclear) in the formulation of scientific work, consequences which we believe have considerable value implications.

Although the target audience for this paper is minerals industry geologists, it should be generally useful to those working in other fields of applied earth science including climatology, oceanography and environmental geochemistry where the required assumptive frameworks are philosophically similar.

If the work of applied geologists in the minerals industry is to be justifiably labelled as scientific we need to pose questions that are, in principle, falsifiable; then we need to subject these appropriately framed questions (hypotheses) to testing (i.e. experiments), and design data collection for these tests in a manner that has the objective of falsifying the hypothesis. We also need to be aware that interpretations of geological data never exist separately to our assumptive framework (i.e. theory).

Why is philosophy of science important for applied geologists?

It is reasonable to ask 'why is following a scientific approach important for an industry or applied

[1]Quantitative Group, PO Box 1304, Fremantle, WA 6959, Australia
[2]School of Earth and Environment, The University of Western Australia, Crawley, WA 6009, Australia
[3]School of Civil Environmental and Mining Engineering, The University of Adelaide, Adelaide, SA 5000, Australia
[4]WH Bryan Mining and Geology Research Centre, University of Queensland

*Corresponding author, email jv@qgroup.net.au

Published by Maney on behalf of the Institute and The AusIMM
Received 19 September 2011; accepted 21 November 2011
DOI 10.1179/1743275811Y.0000000018

geologist?' The answer, we argue, is that minerals industry geologists are practicing scientists. For all scientists the core professional tasks are problem formulation, hypothesis generation, data collection (design and conduct of experiments), interpretation, modelling and prediction based on these models. Following this, ideally, a feedback loop occurs whereby deviations from prediction are used to update or change the hypothesis. These tasks collectively constitute practical implementations of the scientific method.

Expressing this in a more concrete fashion, many of the key activities performed by a mine geologist require us to make predictions based on fragmentary data (for example predicting the location of a coal seam at a point midway between two drill holes in which a coal seam has been recorded). In the mining context, the outcome of these predictions usually has a direct (or indirect) economic consequence. We argue that the quality of prediction is likely to be materially improved if the geologist is firmly and consciously practiced in the scientific method. In addition, if the predictions are framed and presented in terms of the underlying science, the quality of decisions based on those predictions will likewise be improved.

Consequently, an understanding of what science is and how it is done is very important if our work is to be structured and implemented effectively, and – more to the point – if our profession is to sustain and increase our historically critical contributions to the mining industry.

One of the authors has conducted a 'straw poll' of more than 2000 short course participants over the past decade, over 90% of whom were geologists (mostly trained in Australasia, but also in Africa, North and South America, Asia and Europe). Fewer than 5% of these geologists were exposed to philosophy of science during undergraduate university training. Since the great majority of geologists working in the minerals industry evidently received no training in the philosophy of science, we hope that a paper outlining the key aspects of philosophy of science in the context of mine geology will be a practical and helpful contribution. Explicit understanding of the nature of the assumptions employed in designing data collection and doing interpretation and modelling should allow mine geologists to better structure their investigations. This in turn should flow on to development of more robust geological inputs to important (and risky) business decisions.

As an aside, we do not, in this paper, deal specifically with the impact of using probabilistic frameworks as the basis of making predictions or decisions in mine geology (e.g. geostatistics for kriging or simulation models). While this is a very important topic, it must involve detailed consideration of the assumptions specific to statistical models that are beyond the scope of this paper. The interested reader is strongly encouraged to seek out the landmark monograph on this topic by Georges Matheron (1989), which is a comprehensive exposition of the philosophical and practical implications of building probabilistic models for unique phenomena such as mineral deposits.

Philosophy of science

Background

In its broadest sense, science describes the systematic building and organisation of knowledge, in order to explain and predict the world we live in. Philosophy of science is the discipline that examines the system of science itself, considering such aspects as the components of science, methods, assumptions and limitations. This is a broad subject and, in keeping with other branches of philosophy, exact definitions are not readily agreed upon by professional philosophers.

The Oxford Dictionary (http://oxforddictionaries.com/) defines philosophy as '…the study of the fundamental nature of knowledge, reality, and existence'. The two main branches of philosophy are epistemology, the theory of knowledge, especially with regard to its methods, validity, and scope, including the distinction between justified belief and opinion; and metaphysics which deals with the first principles of things, including abstract concepts such as being, knowing, identity, time, and space.

We use the term 'philosophy of science' in this paper as defining the subset of epistemology applied to the activity of science.

The evolving relationship between science and philosophy

All of the sciences (and various branches of mathematics) were originally considered to be parts of philosophy, hence the archaic term for sciences was 'natural philosophy' (Losse, 1980; Rosenberg, 2000). The history of science consists of successive breaking away of individual subject areas (subsets of natural philosophy) to form new disciplines considered distinct and separate. Euclid did this for geometry; Galileo, Kepler and Newton did it for physics; Darwin did it for biology; and Lyell, Murchison and Hutton did it for geology. As an aside, there are many accounts of the origins of the modern science of geology; we recommend Stephen Jay Gould's book (Gould, 1987) on the discovery of 'deep time' as one of the best.

Each time a subject area broke away in the manner described above, the scope of philosophy was reduced with a shrinking set of residual questions left for philosophers (*sensu stricto*). These residual questions, including 'what is the activity called science?' and 'how can we tell whether something is – or is not – scientific?', are core to what is now called 'philosophy of science'. Such questions are important in defining the boundaries of science and to help protect it, for example, when non-scientific activities masquerade as science. In these cases, the fundamental attributes of science need to be understood to identify what is *not* science.

Distinctions between approaches in historical and experimental science

It is evident that scholarship of the philosophy of science did not develop evenly among scientific disciplines. Study of the philosophy of science has been dominated by the experimental sciences, for example physics and chemistry, while the historical sciences (e.g. geology, as well as astronomy, palaeontology and archaeology) have been largely neglected.

Recent work has sought to highlight and describe distinguishing characteristics of the historical sciences, such as the unique nature of the phenomena studied, the temporal separation between initial cause and current state and the interpretation of observations as remnant signs of events and processes (Cleland, 2001, 2002). Such characteristics require modes of reasoning that are in some ways distinctive from those used in experimental sciences (Frodeman, 1995, 2000, 2003). Consideration of these differences is important when deciding how we must design and conduct our investigations and work as mine geologists in order to reasonably claim that our work is scientific.

Traditional concepts of science are largely based upon the modes of investigation followed by physical scientists studying phenomena that can be observed and measured in the laboratory in the present moment, for example in experimental physics, chemistry and biochemistry, to name a few. Experimental physics has been used by Popper (1958) as a type example in his development of arguments, as discussed later in this paper.

There are important differences between the geological sciences and experimental physical sciences which arise from the temporal nature of the phenomena being investigated. Physicists and chemists are generally interested in observing phenomena that can be studied repeatedly, in a controlled environment, in the present moment. In contrast geologists typically investigate unique phenomena that are the products of events that occurred in the past (and often in the very distant past) under uncertain conditions. Present day rocks (and mineral deposits) are the products of historical geological processes superimposed over geological time scales, but the only information now available to us about those processes or events are evidentiary traces contained in the rock (Frodeman, 2000). This imposes some particular restrictions on application of the scientific method to the geological sciences (Cleland, 2001, 2002).

Scientists working in a historical context cannot usually reproduce the events they study in a laboratory today, though some very limited aspects – often under highly constrained conditions and assumptions – may be studied in this way (e.g. the study of melt phase chemistry in experimental petrology; or study of failure modes in shear box experiments). Scientists can, however, look for preserved relicts of features formed by the events that are diagnostic of specific conditions, in what Cleland (2002) calls 'the search for smoking guns'. Sometimes these signs may be clear and unambiguous, for example the presence of a trail of footprints on a bedding layer provides very strong evidence that some creature walked across that exact surface at some unknown time in the past. We are confident to draw this conclusion because modern analogues (footprints in mud left by the passage of a creature) can be observed today. At times we may be happy to extend the observation to make further inferences of varying degrees of certainty – for example, we can be highly certain that at the time the footsteps were imprinted, the horizon lay at the Earth's surface, we can be less certain that the environment was terrestrial and we are even less certain that the creature making the footprints was a particular species of dinosaur.

These sorts of extensions are familiar to geologists, most of whom would have been taught the Principle of Uniformitarianism (Lyell, 1837; Hutton, 1899) and learned that the present is the key to the past. Using modern observable earth processes as an analogy to ascribe similar origin to similar features preserved in the geological record has been invaluable in developing an understanding of the history of our planet. However, in many cases, there is either no modern analogue (e.g. komatiites, flood basalts, asteroid impacts and diatremes) or we simply do not have access to the inferred environment (ophiolites, deep crustal environments, mantle processes). The present is a small and biased subset of all environments that have existed over geological time.

We previously noted that in historical sciences like geology, it is usual to proceed by the examination of evidentiary traces and not direct examination of the actual phenomenon being investigated (Cleland, 2001, 2002). For example, the geologist may be interested in a hypothesis about whether a fault movement pre-dates or post-dates a mineralisation event; however, what are actually investigated are traces of the event, not the event itself. Evidence is thus sought that can distinguish one hypothesis among a set of possible explanations as the best. Such evidence has high value in geology, and the imaginative work (including thought experiments) required is an important (and arguably distinctive) mode of geological scientific thinking. In many cases, evidence that eliminates (or renders less likely) certain hypotheses has at least as much value as evidence that directly supports a hypothesis.

Historical sciences also differ from experimental science in that they are studying *unique occurrences* whereas the latter may observe the same (or essentially the same) phenomena repeatedly. Although there may be families of mineral deposits that share many features, the circumstance of each instance of mineralisation has such a multitude of parameters and boundary conditions that no two deposits can be considered to be identical in the way that two chemical reactions involving the same compounds, run under identical conditions, can. It is impossible that an exact deposit would naturally be created twice. Even so, geologists may be confronted with multiple examples of a given type of geological phenomenon (e.g. many examples of similar fault surfaces, or ancient hydrothermal fluid systems). In such cases, the collective evidence available to geologists resembles a body of evidence akin to that available to an experimental scientist. In such instances, like experimentalists, geologists can formulate generalisations.

These differences between experimental and historical sciences are worth bearing in mind in the discussion that follows.

Scientific models and scientific method

Introduction

The differentiating skill that geologists bring to the minerals industry is their training in geological interpretation – their ability to create coherent explanations (models) from the often sparse observations available to them. Geological interpretations are an essential input to the creation of predictive models necessary for the business of mining, such as block models of metal distribution. Geological interpretations, and the subsequent models we build based on these, are all examples of scientific models.

It is noteworthy that geometric (volume) models of geology have a central place in the minerals industry and their accuracy has huge value implications. When linked with geostatistical block models, they are the fundamental basis of resources and reserves models that are used to design capital developments and expansions, mining schedules, mineral processing strategies and contractually binding product quantity and quality predictions. The financial valuation and evaluation of mineral industry assets is utterly reliant upon sound scientific modelling of key aspects of the exploited materials. Importantly, this modelling is not limited to ore *per se*, but includes marginal and waste materials and their positive and negative impacts on the value chain (Vann *et al.*, 2011).

What is a scientific model?

A scientific model is a set of statements or hypotheses that are postulated to describe the nature of some phenomenon, in order to answer questions about that phenomenon. In mining geology we are typically interested in the profitable extraction of a commodity, for example a metal. In order to plan and execute the mining of this commodity, we require predictive models of the key factors that will influence the value that can be achieved. While concentration (i.e. grade) of the valuable component is clearly a critical variable, geologists know that description of the geometry of important geological features of a deposit underlies the successful spatial prediction of key attributes like grade. Today, such geometric models are usually summarised using computer software as three-dimensional (3D) solid models (wireframes) or surfaces that represent faults, stratigraphic boundaries or alteration fronts. These solid and surface models are scientific models in the sense that they represent or summarise hypotheses about the geometry, extent and character of the geological features. Other examples of scientific models in mine geology include genetic models of mineral deposits, spatial models of grades and other rock properties and models used to aid geotechnical or hydrogeological investigations.

In a general sense, scientific models are representations of hypotheses. These representations may take a number of forms. Whereas they may be purely cognitive (i.e. an unformalised idea), they are generally transferred from cognitive space into a formal (graphical or symbolic) representation.

Symbolic representations may be formal statements of a hypothesis in language for example '…the faulting has normal displacement' or '…the porphyry dykes are the source of the mineralising fluids'. These statements may be very complex, setting out an entire story or scenario to explain the evidence seen by observation and analysis of rocks and their inter-relationships in the present day. Often such explanations require additional supporting symbolic and/or graphical representations. Symbolic representations often include mathematical equations, for example a deterministic formula linking measured rock mass characteristics to likely crushing performance in the plant, or chemical equations specifying alteration or mineralisation chemical reactions and the associated changes in mineralogy (e.g. Best, 2003).

Mapping

In graphical models, the geological hypothesis is often represented as a map, because maps summarise spatial relationships in a succinct way. Maps are literally re-presentation of the cognitive models of the geologist. A map can of course also represent a hypothesis or set of hypotheses in three dimensions. In the simplest case, this may involve cross or longitudinal sections but in more sophisticated cases, 3D models, wireframes or full solid modelling may occur. These days digital wireframe models are commonly used to represent geological surfaces in exploration and resource models; and in many cases when mining industry geologists speak of geological models this is the unspoken implication. However, graphical models can also include two-dimensional maps, diagrams or perspective drawings.

The reduction aspect of geological models

All scientific models have a fundamental *reduction aspect*, i.e. they are simplifications of complex real world situations. This reduction (or reification) is a requirement of all models. The relationship between a scientific model and the phenomena under consideration is analogous to the relationship between a topographic map and the actual topography we are representing. We choose which aspects are important and therefore must be represented in our model. For example, on a topographic map we wish to capture the absolute and relative altitude of the land surface, but not necessarily its permeability or mineral composition.

A consequence of the reduction facet of scientific models is that we must *choose* which aspects of reality need to be represented. We must also choose a scale of resolution and degree of complexity that is to be inherent in the model. The resolution of different features may vary in our model. For example, a spatial model of faults and fractures might be used to aid prediction of the geotechnical stability of an area to be mined. This model will not represent every single discontinuity. However, we may have the objective of capturing as discrete (mapped) surfaces all those discontinuities that are persistent in space over more than 50 m and have other characteristics which indicate they are possible significant failure surfaces. We may choose not to identify minor fractures in a discrete manner, but rather associate estimated fracture frequency values in a block model to specific mapped lithostructural units. This latter decision may be forced upon us by the impossibility of discrete mapping of small fractures using limited data collected from drill holes and existing mine exposures. Hence the resolution of major and minor fractures in the model will be necessarily very different. This means that geology models have a fit-for-purpose aspect and are not useful (and indeed may be misleading) if applied outside their original design limitations. A non-geological analogy is the map of the London underground posted in every station and on trains that many readers will be familiar with (or similar network maps in other cities). These maps have the objective of facilitating navigation and specifically enabling the user to change lines at the right points and know how many stops are left before they should alight from the train. In general, they have no use as a means of answering such questions as 'what is the Euclidean distance travelled between two stations' or 'in which compass direction am I travelling'?

In summary, geological models can only ever be fit-for-purpose within specific limitations. It is critical that geologists be explicit about the limitations of their models when presenting them.

Statistical models in geology

Many numerical or mathematical models in geoscience are necessarily non-deterministic, i.e. statistical models.

This may be either because we have incomplete data (leading to uncertainty), or because there is significant variation in the observed phenomena; usually both these factors apply. Statistical models involve describing the relationship between observed and predicted parameters in stochastic terms and may be as simple as linear regression/correlation models or model fitting (lognormal distributions to identify outliers in geochemical data; for example, Reimann and Filzmoser, 2000). Examples of more complex statistical models include models to assign rock samples into discrete categories based on multi-element geochemical analyses by such methods as Principal Components Analysis or related approaches (Grunsky and Smee, 1999). Statistical models in geology are increasingly based on simulation of some kind. Geostatistical kriged block models (Matheron, 1963) and conditional simulation models (Journel, 1974) are examples of stochastic models now widely used in resource evaluation in the minerals industry.

The systems aspect of geological models

Many models used by applied geologists, for example conceptual exploration models and models for the genesis of mineralisation, are mixtures of cognitive, graphical, chemical, mineralogical, mathematical and statistical components and represent very complex and interdependent systems of hypotheses. Taken *in toto*, such models represent rich descriptions of geological processes and the resultant end products (rocks, alteration assemblages and structures). This raises the important issue that models of complicated phenomena like geology have a systems aspect which cannot be ignored. A profound or deep understanding of complex situations nearly always requires that we not only acknowledge the variability and uncertainty in the system, but that we also explicitly model the systems themselves (Deming, 1986). Modelling of isolated parts of a system without an understanding of the interconnections and linkages is fraught with danger.

Parsimony

The principle of parsimony should be adhered to when building models. In short, this implies that the objective is to build models that have the simplest possible combination of attributes, whilst remaining fit-for-purpose (Matheron, 1989). The key issue here is that simplification is necessary but over-simplification will result in inadequate predictive power. For example, simplifying spatial models of grades by averaging incurs the risk of inadequate prediction for any situation where grade variability is important.

Predictive versus explanatory modes

Scientific models have two main modes of use. The first is to explain existing observations, for example, 'why do we always see carbonate rocks capping a certain sequence of marine sediments'? The second is to predict new observations and this is usually very important for the applied geologist working in the minerals industry. For example, the geologist wants to predict such things as:

- Is the mineralisation revealed in two drill intercepts separated by 50 m likely to be connected in space, i.e. is the mineralisation continuous between the two holes?
- If we drill a hole beneath the currently drilled part of the mineralisation, where should we drill it to maximise the chances of intersecting strong mineralisation?
- Are the grades likely to increase or decrease significantly with depth?
- What will the copper grade of a certain block of ground be when we extract it in the mining sequence?
- How far will the down-thrown part of the ore zone be offset by a particular fault?

Scientific method

Scientific method is a systematic approach to investigation of phenomena, characterised by observation and measurement of some aspects of the phenomenon, generation of statements (hypotheses) from which predictions may be logically extended, and design of experiments to test the strength of such predictions. Depending on the outcomes of testing, hypotheses may be rejected, strengthened or modified. The scientific method is thus embedded in the generation and testing of hypotheses against observed results. Hypotheses may take the form of equations, verbal statements or other models or representations of reality, such as maps and wire-framed solids.

Historically the usefulness of a scientific model depended on its success in predicting the outcomes of experiments or new observations. The evolution and refinement of a scientific model proceeds by making predictions, based on a model or set of hypotheses, and then comparing the outcomes of experiments or observations to those predicted by the model. A model is therefore always *interim:* it will be refined or even abandoned if it fails to predict existing or new observations.

Popper and falsification

The idea of falsification

The philosopher Karl Popper proposed that the fundamental attribute of scientific models is that they be falsifiable and argued that the essential mechanism for testing and refinement of scientific models is to attempt falsification (Popper, 1958, 1963). He listed the salient aspects that constitute a scientific theory, which we paraphrase as follows.

Good scientific theories have embedded prohibitions: they forbid certain things to happen. The more a theory forbids, the better it is. Furthermore, a theory which is not refutable by any conceivable event or test is non-scientific; a scientific theory must be refutable by testing (this is the definition of falsification). Every genuine test of a theory is thus an attempt to falsify it, or to refute it; i.e. a theory is scientific only to the extent that it is falsifiable (testable in a way which can yield refutation).

We agree that falsifiability is a key objective for geologists wishing to claim scientific status for their models; in other words our task is to constructively criticise our models by designing tests that maximise the likelihood of disproving our geological hypotheses. There is an absolutely essential prior step, which is ensuring that the purpose of the model is properly defined and understood. If a hypothesis withstands such attempts at falsification, it is considered to be more reliable, but still not proved. If it repeatedly withtands such tests, we can increase our confidence in the hypothesis. The usefulness of the model increases in the sense that it is a source of more and more reliable predictions that are consistent with new observations or experimental results.

Scientific models in applied geology therefore *must* be based on hypotheses that are, at least in principle, falsifiable, and it can be further argued that only hypotheses that are falsifiable are valid. It is not possible to advance knowledge if a theory cannot be tested. For example, the hypothesis that a supernatural being created the universe by laying an egg cannot be subjected to a falsification test and resides in the realm of theology, not science. The hypotheses that constitute astrology are also mostly untestable and cannot be compared to those in astrophysics.

The concept of falsifiability has practical implications for applied geologists: we must endeavour to frame all our hypotheses in terms that are, in principle, falsifiable. For example, it is possible to test the hypothesis that ore is present at a given location by conducting an experiment (drilling a hole at that location). We should design our tests (for example, drilling) to be attempts at falsifying a well constructed hypothesis. Such drilling will be most useful for improving and refining our model if a conscious strategy of designing holes to challenge or invalidate our model is adopted. The same is also true if specific holes are sited to test areas of disagreement between two (or more) plausible models. The challenge lies in identifying those elements of difference between competing models that would result in material differences in economic outcome. To do this requires consideration of the full space of plausible hypotheses, and as well some subjective assessment of the likelihood of occurrence of those hypotheses.

In a hierarchy of importance for the activities of a mine geologist, we would place interpretation (i.e. geological modelling) at the highest level. Data acquisition is essentially mechanical (although still governed by theory). Spatial model construction (grade or other rock attributes) is largely concerned with the science of uncertainty. It is interpretation that is the uniquely geological scientific activity for mine geologists – and it is here that the philosophical framework strays most from a strict Popperian view of the world.

A more sophisticated idea of falsification

Many people, Popper included, realised that problems are encountered when attempting to move from the logic of falsifiability to the real-world practice of testing, i.e. attempted falsification (Champion, 2011). A strict adherence to the necessity for falsification has been referred to as falsificationism (Feyerabend, 1993). The real value of falsification to the practice of science is that the next stage of progress is embedded in the last because a solution includes testing. Falsification is very practically useful when applied to constructing and testing geometric models (usually as wireframes of interpreted lithology, alteration, weathering and structures) used for resource evaluation. In our opinion interpretations can be efficiently improved using a falsification framework and we therefore regard it as a pragmatic and necessary tool for mine geologists.

A valuable feature of falsification is that it compliments other modes of scientific investigation. Investigation of a single hypothesis in falsification mode is weak compared to using a multiplicity of inconsistent theories (Feyerabend, 1993). In fact such an approach of using multiple working hypotheses was originally proposed by the geologist T. C. Chamberlin in 1897, and we return to this shortly.

There is an ambiguity associated with falsification because in every instance a negative result may be attributed to either a flawed hypothesis or an error or inaccuracy in the experiment, observation or underlying assumptions (Feyerabend, 1993). The possibility of flawed underlying assumptions must always be considered. On the other hand, as Cleland (2001) noted, students are usually acutely aware that procedural errors are a possible cause of a failed experiment. This is because repetition of classical experiments during laboratory exercises does not necessarily replicate the well-established result, e.g. due to malfunctions or contamination of equipment. It is similarly possible that we get false signals in geological investigations. In summary, failing a test does not necessarily disprove the hypothesis and careful consideration of the context of the experiment is always important.

A classic example of the use of incorrect theoretical assumptions is the Copernican hypothesis for a heliocentric solar system, in which the Earth orbits the Sun, which was widely rejected by contemporaries of Copernicus (Chalmers, 1999). One falsifying experiment proposed to test the hypothesis posited that the stars should exhibit measureable parallax error when the Earth was situated at opposite sides of the sun. Measurements were duly made and no such parallax could be observed. The assumptive flaw was, of course, that the stars were close enough for such a difference to be measureable – which they are not.

A significant criticism of Popper's theories was proposed by Thomas Kuhn (1996, first published 1970), who pointed out that scientists almost never practise strict falsification. In fact, if a prediction fails a test, scientists often, quite rationally, engage in a search for conditions that might be responsible other than those proposed by the hypothesis. This amounts to a relaxation or adjustment or re-framing of the hypothesis. This has been described as exercising the option of salvaging a hypothesis by rejecting some component assumption (Cleland, 2001). In other words, falsification may involve modification of the hypothesis by rejecting some auxiliary assumption rather than the total rejection implicit by strict adherence to falsification.

In the case of mine geology, this raises the important issue that the test of the hypothesis is always conceived within a conceptual framework which could itself be flawed; and with experimental approaches that can be erroneous. We consider that a modified falsification approach is usually advisable in geological investigations with the following characteristics:

(i) The hypotheses must be stated so that they are falsifiable
(ii) After failure of a test other possibilities should be investigated before rejection of the hypothesis, which may include, in particular:
   - the possibility of error in the test (as previously discussed)
   - the hypothesis itself should be evaluated for assumptions that may be erroneous or flawed.

For novel ideas, the geologists often must go looking for confirmatory evidence to convince themselves that the concept being investigated is fertile. This has been described as '…the search for a smoking gun [which] is a search for supporting evidence for a hypothesis' (Cleland, 2002, p. 483). A classic geological example of the search for a smoking gun is the hypothesis that an asteroid impact was responsible for the extinction of the dinosaurs and many other taxa at the Cretaceous–Tertiary (KT) boundary (Alvarez and Vann, 1978; and Alvarez *et al.*, 1980). In this instance, the primary hypothesis (that there was a major impact at the KT boundary) suggested looking for other evidence of impact, such as elevated Iridium and Osmium levels and presence of shocked quartz in the KT boundary sedimentary rocks. This evidence was sought and found, lending more weight to the hypothesis.

Geologists can and must look for confirmatory evidence for novel hypotheses. This approach requires that they need to intelligently visualise what types of evidence must be sought. This mode of investigative working is similar to that conducted at a crime scene. Such evidence is often well hidden in the complex, messy, partially preserved and incompletely exposed geological record. If found, the smoking gun does not prove our hypothesis but such evidence will usually significantly strengthen our confidence in a hypothesis, especially if we specify the type of evidence expected before we go looking for it. If we predict such evidence and cannot find it, our hypothesis is unconfirmed. Over time repeated failure to find confirmatory evidence will erode our confidence in the hypothesis, which in the framework of Chamberlin's (1897) multiple working hypotheses theory, may lead us to prefer alternative hypotheses.

Kuhn and paradigm shift

Kuhn (1996) concluded, after a study of the history of various scientific developments, that major advance in science have occurred not by sequential falsification, but via successive paradigms, shattered by revolutions in which these paradigms were overturned. Kuhn also discussed the important role that sociology plays in the beliefs and behaviours of scientific communities. The interested reader is encouraged to read Kuhn (1996) along with the excellent summary and critical analysis of Kuhn's contributions by Chalmers (1999). We will now summarise this philosophy using examples drawn from geology.

A mature science is always characterised by a single paradigm (Kuhn 1996). Although Kuhn does not explicitly define mature, we take it to mean that a discipline has built up a significant body of observations and constructed robust theoretical frameworks that accommodate these observations. Others (e.g. Burroughs, 2008) have argued that maturity requires an emergence of informed critiques, focused on the limits of methods of analysis employed in a given discipline.

Similarly the idea of a paradigm is not succinctly defined by Kuhn, but we take it to be the sum of the general theoretical assumptions and laws, along with techniques for their application, that the members of a particular scientific community adopt (Chalmers, 1999). In Kuhn's terminology, scientists working within a given paradigm practise what he calls normal science. If new observations arise that seriously challenge the existing paradigm, or if a new theoretical framework is proposed that seems to better explain certain observations (or is a better basis for prediction), a crisis state occurs. Such a crisis is resolved by the emergence of a new paradigm, in a discontinuous change referred to by Kuhn (1996) as a scientific revolution.

A classical example of a paradigm shift is the change from Newtonian to quantum/relativity physics, but subdisciplines can also undergo such revolutions. One example used by Kuhn (1996) of a revolution in a subdiscipline is the emergence of modern ideas about electric currents. The plate tectonic revolution is a relevant example for geology.

Geological examples

A comprehensive discussion of the history of ideas and people behind the plate tectonic revolution is given by Oreskes and Le Grande (2001). An in-depth discussion of the philosophical implications and some interesting observations on the sociological and psychological dimensions of the plate tectonic revolution is given by Solomon (1992). Here we present an abbreviated account and draw some general conclusions.

Prior to the emergence of new types of data, including detailed seafloor bathymetry and sea floor magnetic imagery, the prevailing geological paradigm was of an immobile earth in which the continents and oceans occupied essentially unchanging positions. Previous contentions about mobile continents from the time of Wegner (1924) and Holmes (1929) were initially rejected by the overwhelming majority. There was consensus that the continents did not drift, collide or break apart. It is sometimes argued that this rejection of drift was based on lack of mechanisms for such large scale crustal mobility, but this is not correct. There was a wide debate on possible mechanisms for continental drift (Oreskes, 2001), and in fact Holmes (1929) laid out a remarkable account that foreshadowed plate tectonics, at least in a cursory form. The work of Holmes (1929) even discusses the hitherto unknown concept of subduction.

It was not until the 1960s that new observations by geophysicists of symmetric sea floor striping started the shift towards acceptance of continental drift (Oreskes and Le Grande, 2001; and references therein). There were considerable sociological impediments to acceptance of this idea (Solomon, 1992). The paradigm shift was more rapid among some communities in geology than others (e.g. acceptance was much slower in North America). Once the paradigm shift was underway, the hypothesis of plate tectonics spawned a range of testable predictions, some entirely new, which had profound impacts in economic geology. A survey of ore deposit models will show that the plate tectonic context is now a critical component of such models and has had considerable predictive success.

In summary, overcoming the status quo is usually difficult, and does not proceeded by simple falsification pathways. In the case of plate tectonics, many of the essential ideas existed long before the paradigm shift, but were rejected by the mainstream (ie, so-called normal science). To Kuhn a paradigm is a theoretical framework or structure that becomes the boundary conditions of thought and action within a given scientific field. Scientists find it hard to reason outside of such existent theoretical frameworks. Kuhn argued that scientists must, at least to some degree, be uncritical of the paradigm in order to be able to investigate detailed aspects of that paradigm. It is therefore often very difficult for scientists to let go of long held views even when the evidence seems clear that a new framework explains things much better.

In the case of geological interpretations and ideas at deposit level, we have seen numerous examples of deeply held beliefs around specific ideas by geologists, where conflicting evidence is resisted strongly. This is one reason why, as much as familiarity with a deposit is valuable, a well-reasoned and constructive challenge to the status quo may be more so. These challenges often come from those with less commitment to previous ideas (Solomon, 1992).

An important conclusion for mine geologists to draw from Kuhn (1996) is that there are no pure facts. Theory, or more broadly, paradigm, is always a framework for any observations. Whereas the framework enables hypothesis construction, it can also constrain our thinking as geologists. Once we are anchored within a given paradigm, identifying potentially contradictory or falsifying observations becomes harder. The possibility that we fail to see or we misinterpret evidence because of this anchoring is heightened for a range of socio-cognitive reasons as explored by Solomon (1992).

A good example of a paradigm shift relating to a specific deposit model is the world class Olympic Dam Cu–U–Au deposit in South Australia. Original exploration models and early publications emphasised a syngenetic or syn-diagenetic model for deposit formation in a sedimentary breccia (Roberts and Hudson, 1983). Continuing data collection as the deposit was further explored and then accessed by mine workings resulted in a major re-evaluation of the deposit origin (Selby, 1991), leading to the current hydrothermal breccia model.

Multiple working hypotheses

We have argued above that a key to development of robust models in mine geology is that they be examined critically in the spirit of falsification. It is advisable for scientists to keep more than one competing hypothesis alive (Chamberlin, 1897; Feyerabend, 1993). Such models should be mutually contradictory, whilst agreeing with the available data. This idea has great power in mine geology, where even relatively minor changes in interpretation may have serious economic implications, and major differences may have economically disastrous consequences. This is true for both the geological interpretation (for instance in interpreting shear zone continuity between *these* two logged shear intersections, rather than *those* two), and for the translation of these interpretations into 3D wireframes (for example, connecting two contact points in adjacent holes with a straight line versus inserting additional control points if the contact is interpreted to be curved).

There are clearly major benefits in being able to assess the economic impact of alternative hypotheses, and thus justify the expenditure necessary to test (or attempt to falsify) these hypotheses. In the first instance it is necessary to identify the key assumptions, then envisage plausible alternatives and test these by directed data acquisition (e.g. drill holes in strategic locations where conflicting hypotheses predict different geometry). Until recently, it has been difficult (or impractical) to generate and evaluate even limited numbers of alternative geological models (or more correctly the 3D computer representations of these). Increasingly, though, the generation of multiple, divergent, digital 3D models is practically achievable because of faster computers and improved automated or semi-automated 3D modelling tools. A straightforward and insightful example of the use of multiple geological interpretations in resource risk analysis is given by Jackson *et al.* (2003).

If models are to be evaluated rigorously then having external critical review, as well as a robust internal critical review culture, is essential. We believe that the establishment of multiple interpretive teams for major capital projects is a prudent and practical risk reduction (and thus value creating) mechanism, although this process must be well managed. It is difficult for an individual team (and more so an individual geologist) to develop a positive and genuinely critical environment for the generation of geological models. The interpretive process requires that we invest effort in imagining ideas, and it is human nature that once we have invested that energy we become the champion of those ideas. Geologists are not exempt: T. C. Chamberlain, a nineteenth century geologist, described this eloquently:

> *'The moment one has offered an original explanation for a phenomenon which seems satisfactory, that moment affection for his intellectual child springs into existence; and as the explanation grows into a definite theory, his parental affections cluster about his offspring and it grows more dear to him. While he persuades himself that he holds it still as tentative, it is none the less lovingly tentative and not impartially and intemperately tentative' (Chamberlin, 1897, p. 358).*

It is interesting that the ideas of Chamberlin, whilst known by a small proportion of mine geologists, have wide currency and use in other fields (for example in biology, see Platt 1964; Elliot and Brook, 2007).

The problem of model validation

Verification or validation of scientific models of complex natural systems is impossible according to Oreskes *et al.* (1994) and Oreskes (1998). This assertion has direct relevance for numerical models like resource estimates and conditional simulation models of spatial variables (Journel, 1974), which are increasingly used in mining applications.

In essence, agreement between models and new observations or predictions can only be taken as partial confirmation since acquisition of further data may yet invalidate the model. This is similar to the point made by Cleland (2001, 2002) that predicting and then confirming a 'smoking gun' increases the confidence we have in a model but that such confirmation is always interim and partial (and not without the dangers alluded to by Popper). The incomplete access we have to natural phenomena (we never have full knowledge of the orebody at every scale) means that models can only be evaluated and deemed to be fit-for-purpose and cannot be validated in the strict sense. This is an important practical point because geologists must communicate the uncertainties in their models clearly in order to justify improvements. Giving geological models the status of truth is always a mistake. Operating in a mode of multiple working hypotheses is a counter weight to this tendency to become too wedded to a single model.

It is also arguable that the primary value of models is heuristic, i.e. to be used in a pragmatic way to guide decisions and further investigations (Oreskes *et al.*, 1994). This aligns with the statement by George Box quoted at the opening of this paper that the model is by definition wrong, at least to some degree, because it is a model not reality. The real question is whether the model is useful – can it practically help guide better decisions?

The use of so called heuristic models (experience based models, akin to rules-of-thumb) as a basis for decision making where there is uncertainty has been investigated in many fields (Tversky and Kahneman, 1974; Kahneman *et al.*, 1982). The influence of psychology in making decisions in the face of the inherent uncertainty in geology is an important area for current and future research, for example there is ongoing research into this area of behavioural geoscience at the Centre for Exploration Targeting (University of Western Australia and Curtin University of Technology).

Rejoinder: the truth of climate models

As an aside, those discussing climate science and climate models (which are numerical models of complex natural systems) would do well to heed the issues raised above relating to the status of models.

Statements that numerical climate models (or predictions based on them) have the status of truth; and that, consequently, debate about the validity of predictions from these models is finished, are highly misleading. It is a fundamental attribute of predictive models of complex natural systems (like mineral deposits or climate systems) that they cannot be verified; such models have intrinsically interim status. Successful comparison of predictions generated by such models against new observations can be confirmatory and thus increase confidence in the hypotheses encapsulated in the model. Such agreements do not lend to the model the status of truth; however, or end the scientific debate.

It is true that some hypotheses, especially those that can be subjected to repeated controlled experiment, have been subjected to such repetitive scrutiny that it is very hard to imagine them being overturned. The basic laws of motion in the physics of the macroscopic universe fall firmly into this category.

The idea of retrodiction in geology, meaning that a hypothesis (or hypotheses) can be framed and then the past record repeatedly interrogated to look for confirmatory traces, has been proposed (Kitts, 1978). The idea of the biological evolution of taxa over geological time is an example of a hypothesis that has been confirmed by repeated evidence in this retrodictive mode (Dawkins, 2009). In legal parlance, it is beyond reasonable doubt.

Most predictive models do not have this status in geology (or climate science) and remain interim in nature, even if we steadily acquire more confidence in them as we fail to falsify them, or gather more confirmatory evidence. It serves well for mine geologists to remember this, and communicate it in a business context.

Conclusions

Constructing testable models to explain reality is the definitive aspect of any activity claiming to be scientific. In mine geology, for example, our primary job is to build models which accurately predict reality to an acceptable degree, be they geological models, grade control models or resource models. In mine geology we are often in the excellent position of having hypothesis-testing options, such as additional samples or new mine openings that can provide relatively rapid feedback on how good our predictive models are.

In applied geology, the idea of falsification is very practical and useful but needs to be considered in a sophisticated way. In particular, the collection of data and design of experiments must ensure that the hypotheses we frame are falsifiable, at least in principle. If not, our work cannot be defended as being scientific. Note that some theories may be falsifiable in principle but not in practice using current technology and methods (Einstein's famous thought experiments regarding quantum mechanics come to mind).

Another important framework for considering the work of applied geologists is the idea of seeking confirmatory evidence (especially in the case of novel ideas). If we can find evidentiary traces that were predicted from a hypothesis prior to examination of the geological record, we increase the confidence we have in a model significantly. Testing of models with more than a single explanatory hypothesis is particularly powerful in the case of retrodictive modes of science, where we are trying to explain a set of evidence present today (in the geological record for example) that could arise from multiple possible mechanisms in the past. This drives the usefulness of multiple working hypotheses as a mode of thinking in geology. The method of multiple working hypotheses is a valuable and practical means of insurance against many of the problems associated with application of scientific method for historical sciences. In fact, we would argue that this is the major conclusion to take away from this paper regarding the improvement of science in applied geology.

The framework of assumptions that geologists use is often unchallenged. The idea of working within an unchallenged paradigm is not necessarily negative, and may well be required to generate useful results. However; it is important to be attentive to (and on the lookout for) constructive challenges to the status quo, because from these come all really important new scientific breakthroughs.

Acknowledgements

Much of the material in this paper was originally published in a different form as Vann and Stewart (2011) and presented to an AusIMM conference of about 300 professional mining geologists in Queenstown, New Zealand. This significantly revised version has been re-structured and rewritten to make it suitable to a more general audience and as well as having the addition of appreciably expanded material on scientific models and method.

The ideas expressed here have evolved through discussions with many of our industry colleagues and participants in mining geology related training courses run by both authors. Toby Davis of Impel Geoscience and Stephen Coward at QG also thanked for useful philosophical discussions over recent years. Toby Davis is also acknowledged for his thoughtful review of an early version of the manuscript, which along with reviews of later versions by Neil Phillips of the University of Melbourne and Phillipsgold Pty Ltd, and Simon Dominy of Snowden Mining Industry Consultants and the University of Ballarat improved the paper. Ruth Boardman of QG and an anonymous reviewer of the paper submitted to this journal prompted the abovementioned restructuring. Any remaining deficiencies are – of course – entirely the responsibility of the authors.

The first author would finally like to personally acknowledge his mentor in geology and geostatistics, Pedro Carrasco of Codelco, who sadly passed away in 2010. Pedro often pointed to the importance of philosophical aspects of geology and geostatistics in mining and this subject remained a passion throughout his life. We urge all applied geologists to be philosophical in their work.

References

Alvarez, L. W., Alvarez, L., Asaro, F. and Michel, H V. 1980. Extraterrestrial cause for the Cretaceous-Tertiary extinction, *Science*, **208**, 1095–1108.

Alvarez, L. W. and Vann, D. W. 1978. Comments and replies on 'Biostratigraphy and magnetostratigraphy of Paleocene terrestrial deposits, San Juan Basin, New Mexico', *Geology*, **7**, (1), 66–67.

Best, M. G. 2003. Igneous and metamorphic petrology, 2nd edn, 752, New York, Wiley Blackwell.

Burroughs, W. J. 2008. Climate change in prehistory: the end of the reign of chaos, 372, Cambridge, Cambridge University Press.

Box, G. E. P. and Draper N. R. 1987. Empirical model-building and response surfaces, Wiley Series in Probability and Statistics, 688, New York, John Wiley.

Chalmers, A. F. 1999. What is this thing called science?, 3rd edn, 288, Hackett, University of Queensland Press.

Chamberlin, T. C. 1897. The method of multiple working hypothesis, *The Scientific Monthly*, 59, 356–362. (Reprinted from *The Journal of Geology*, **5**, 837–848).

Champion, R. 2011. http://www.the-rathouse.com/shortreviews/WhatisThisThingCalledScience.html (accessed 3 June 2011).

Cleland, C. E. 2001. Historical science, experimental science and the scientific method, *Geology*, **29**, 987–990.

Cleland, C. E. 2002. Methodological and epistemic differences between historical science and experimental science, *Philosophy of Science*, **69**, 474–496.

Dawkins, R. 2009. The greatest show on Earth: the evidence for evolution, 480, New York, Free Press – Simon and Schuster.

Deming, W. E. 1986. Out of the crisis, 507, Cambridge, Cambridge University Press.

Elliot, L. P. and Brook, B. W. 2007. Revisiting Chamberlin: multiple working hypotheses for the 21st Century, *BioScience*, **57**, 608–614.

Feyerabend, P. 1993. Against method, 3rd edn, 279, London, Verso.

Frodeman, R. 1995. Geological reasoning: geology as an interpretive and historical science, *Geol. Soc. Am. Bull.*, **107**, 960–968.

Frodeman, R. 2000. Earth matters: the earth sciences, philosophy, and the claims of community, 209, Upper Saddle River, NJ, Prentice-Hall.

Frodeman, R. 2003. Geo-logic: breaking ground between philosophy and the earth sciences, 137, New York, SUNY Press.

Grunsky, E. C. and Smee, B. W. 1999. The differentiation of soil types and mineralization from multi-element geochemistry using multivariate methods and digital topography, *J. Geochem. Explor.*, **67**, 287–299.

Gould, S. J. 1987. Time's arrow, time's cycle: myth and metaphor in the discovery of geological time, 252, Cambridge, MA, Harvard University Press.

Holmes, A. 1929. A review of the continental drift hypothesis, *Min. Mag.*, **1**, 2–17.

Hutton, J. 1899. Theory of the earth (4 Volumes), London, The Geological Society.

Jackson, S., Fredericksen, D., Stewart, M., Vann, J., Burke, A., Dugdale, J. and Bertoli, O. 2003. Geological and geostatistical risk at the Golden Gift and Magdala gold deposits, Stawell, Victoria, Australia, Proc. 5th In. Min. Geol. Conf., 215–223, Melbourne, The Australasian Institute of Mining and Metallurgy.

Journel, A. G. 1974. Geostatistics for conditional simulation of ore bodies, *Econ. Geol.*, **69**, 673–697.

Kitts, D. B. 1978. Retrodiction in geology, Proceedings of the Biennial Meeting of the Philosophy of Science Association, Vol. 2: Symposia and Invited Papers, 215–226, Chicago, IL, The University of Chicago Press.

Kahneman, D., Slovic, P. and Tversky, A. (eds.) 1982. Judgment under uncertainty: heuristics and biases, 555, Cambridge, Cambridge University Press.

Kuhn, T. 1996. The structure of scientific revolutions, 3rd edn, 212, Chicago, IL, University of Chicago Press.

Losse, J. 1980. A historical introduction to the philosophy of science, 2nd edn, 248, Oxford, Oxford University Press.

Lyell, C. 1837. Principles of geology, (2 Volumes), 1st American edn, Pittsburg, PA, John L. Kay & Co.

Matheron, G. 1963. Principles of geostatistics, *Econ. Geol.*, **58**, 1246–1266.

Matheron, G. 1989. Estimating and choosing, 141, Berlin, Springer-Verlag.

Oreskes, N. 1998. Evaluation (not validation) of quantitative models, *Environ. Health Perspect.*, **106**, (Supplement 6), 1453–1460.

Oreskes, N. 2001. From continental drift to plate tectonics, in Plate tectonics: an insiders history of the modern theory of the earth, (ed. N. Oreskes and H. Le, Grande H.), 3–17, Cambridge, MA, Westview.

Oreskes, N. and Le Grande, H. 2001. Plate tectonics: an insiders history of the modern theory of the earth, 424, Cambridge, MA, Westview.

Oreskes, N., Shrader-Frechette, K. and Belitz, K. 1994. Verification, validation, and confirmation of numerical models in the earth sciences, *Science*, **263**, 641–646.

Platt, J. 1964. Strong inference, *Science*, **146**, 347–353.

Popper, K. 1958. The logic of scientific discovery, 480, New York, Harper and Row.

Popper, K. 1963. Conjectures and refutations, 582, London, Routledge.

Reimann, C. and Filzmoser, P. 2000. Normal and lognormal data distribution in geochemistry: death of a myth. Consequences for the statistical treatment of geochemical and environmental data, *Environ. Geol.*, **39**, (9), 1001–1004.

Roberts, D. E. and Hudson, G. R. T. 1983. The Olympic Dam copper-uranium-gold deposit, Roxby Downs, South Australia, *Econ. Geol.*, **78**, 799–822.

Rosenberg, A. 2000. Philosophy of science: a contemporary introduction, 191, London, Routledge.

Selby, J. 1991. Olympic Dam: a giant Australian ore deposit, *Geol. Today*, **7**, 24–27.

Solomon, M. 1992. Scientific rationality and human reasoning, *Philos. Sci.*, **59**, 439–455.

Tversky, A. and Kahneman, D. 1974. Judgment under uncertainty: heuristics and biases, *Science*, **185**, (4157), 1124–1131.

Vann, J., Jackson, J., Coward, S. and Dunham, S. 2011. The Geomet curve – a model for implementation of geometallurgy, in Proc. 1st AusIMM Int. Geometall. Conf., 35–43, Brisbane, AusIMM.

Vann, J. and Stewart, M. 2011. I think, therefore I add value: philosophy of science as a practical tool for the mine geologist, in Proc. 8th Int. Mine Geol. Conf., 49–58, Queenstown, AusIMM.

Wegner, A. 1924. The origin of continents and oceans, (Excerpt published in Foundations of biogeography – classic papers with commentaries, (ed. Lomolino *et al.*), 2004, 227–294, Chicago, IL, University of Chicago Press.