# Hotspots in Hindsight

Bruce R. Julian, Gillian Foulger
Dept. of Earth Sciences, Durham University, UK

Oliver Hatfield, Samuel E. Jackson, Emma Simpson, Jochen Einbeck
Dept. of Mathematics, Durham University, UK

Andrew Moore
James Cook University, Townsville QLD 4811 Australia
and
Dept. of Geology, Rhodes University, Grahamstown, South Africa

*The first principle is that you must not fool yourself – and you are the easiest person to fool.* — Richard P. Feynman, 1974

## ABSTRACT

Thorne et al. (2004), Torsvik et al. (2010; 2006) and Burke et al. (2008) have suggested that the locations of melting anomalies ("hot spots") and the original locations of large igneous provinces ("LIPs") and kimberlite pipes, lie preferentially above the margins of two "large lower-mantle shear velocity provinces", or LLSVPs, near the bottom of the mantle, and that the geographical correlations have high confidence levels (> 99.9999%) (Burke et al., 2008, Fig. 5). They conclude that the LLSVP margins are "Plume-Generation Zones", and that deep-mantle plumes cause hot spots, LIPs, and kimberlites. This conclusion raises questions about what physical processes could be responsible, because, for example, the LLSVPs are apparently dense and not abnormally hot (Trampert et al., 2004).

The supposed LIP-hot spot-LLSVP correlations probably are examples of the "Hindsight Heresy" (Acton, 1959), of performing a statistical test using the same data sample that led to the initial formulation of a hypothesis. In this process, an analyst will consider and reject many competing hypotheses, but will not adjust statistical assessments correspondingly. Furthermore, an analyst will test extreme deviations of the data, , but not take this fact into account. "Hindsight heresy" errors are particularly problematical in Earth science, where it often is impossible to conduct controlled experiments.

For random locations on the globe, the number of points within a specified distance of a given curve follows a cumulative binomial distribution. We use this fact to test the statistical significance of the observed hot spot-LLSVP correlation using several hot-spot catalogs and mantle models. The results indicate that the actual confidence levels of the correlations are two or three orders of magnitude smaller than claimed. The tests also show that hot spots correlate well with presumably shallowly rooted features such as spreading plate boundaries.

Nevertheless, the correlations are significant at confidence levels in excess of 99%. But this is confidence that the null hypothesis of random coincidence is wrong. It is not confidence about what hypothesis is correct. The correlations probably are symptoms of as-yet-unidentified processes.

These observations emphasize importance of the distinction between correlation and causation and underline the necessity of taking geological factors into account. Consideration of the kimberlite dataset in the context of geological setting, for example, suggests that the apparent association with the LLSVP margins results from the fact that the Kaapvaal craton, the site of most of the kimberlites considered, lies in southern Africa, and that kimberlite eruptions are sensitive primarily to stress in the lithosphere.

## INTRODUCTION

Following a suggestion of Thorne et al. (2004), Torsvik et al. (2010; 2006) and Burke et al. (2008) have suggested that the locations of present-day melting anomalies ("Hot Spots"), and also the original locations of eruption of flood basalts ("Large Igneous Provinces" or LIPs) and kimberlites, correlate closely with the margins of two large anomalies in the seismic shear-wave speed $V_S$ in the deep mantle ("large lower-mantle shear velocity provinces", or LLSVPs) (Figure 1). They claim a high level of statistical confidence for these correlations, with the probability of the LIP-LLSVP agreement, for example, resulting from chance being less than $10^{-6}$ (Burke et al., 2008, Fig, 5), and infer that thermal plumes rising from "Plume Generation Zones", or PGZs, at the margins of the LLSVPs, cause hot spots, LIPs, and kimberlites and have persisted for at least hundreds of millions of years (Burke et al., 2008). Several things about these inferences, however, suggest that caution is warranted.

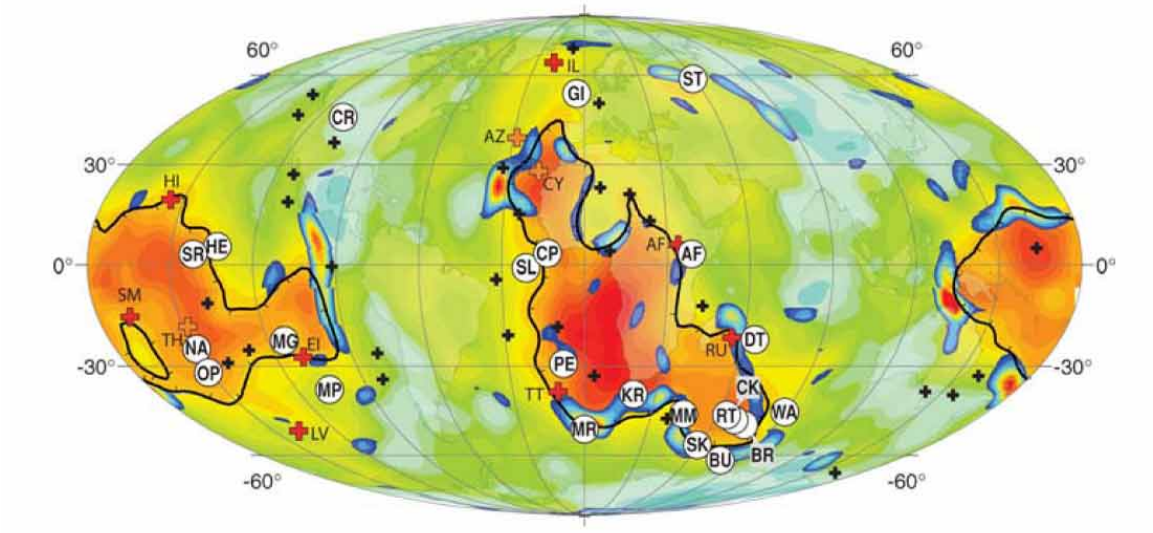


Figure 1: Hot spots (crosses) and inferred original LIP eruption sites (labeled circles), superimposed on colors representing variations $\Delta V_S$ in the shear-wave speed from its average value at 2800 km depth in the model SMEAN of Becker and Boschi (2002), from Figure 10 of Torsvik *et al.* (2006). The central meridian has longitude zero and the black curve is the −1% $\Delta V_S$ contour.

The confidence levels of the proposed correlations can be tested only using the same data that originally led to the formulation of the hypotheses. The hypotheses will thus necessarily match the data well, but statistical tests will use null hypotheses based on assumptions of independence.

Comparisons between LLSVPs and hot spots, kimberlites, or LIPs involve many choices that could affect the result:

• There are many published three-dimensional models of the mantle, differing significantly from one another. Burke and Torsvik (2004) initially used SMEAN (Becker and Boschi, 2002), an average of three different models (and therefore not necessarily consistent with seismic data), but in later work included models of Kuo et al. (2000) and Castle et al. (2000) because they fit the hypothesis better. These models contain small anomalies separated from the main LLSVPs that are closer to the inferred eruption locations of the Siberian Traps and the Columbia Plateau Basalts.

• There is a choice of what physical variable to use (compressional-wave speed, shear-wave speed, bulk sound speed, density, anelasticity, …), and what kind of feature to use (high values, low values, average values, intermediate values, high gradients, …). In this menu, the $-1\%\ \Delta V_S$ contour is an "intermediate value" option, one of perhaps 20 alternatives.

• Hundreds of volcanic areas have been labeled as hot spots, and there are many different published catalogs of them. Torsvik et al. (2006) used a catalog assembled by a co-investigator, Steinberger (2000). The same kinds of choices arise for LIPs and kimberlites. How sensitive are the correlations with LLSVPs to these choices?

• To compare LIPs or kimberlites with the present-day LLSVPs, one must restore them to their original eruption sites, correcting for plate motions over an interval of up to 250 million years. Relative plate motions probably are fairly well known for the Mesozoic and Cenozoic, but the absolute motions are subject to significant uncertainties, particularly in longitude. Quantifying plate motion (or lack thereof) with respect to the deep mantle requires several highly speculative assumptions.

• One would not expect surface manifestations such as hot spots or LIPs to lie directly above plume-generation zones. It is physically impossible for mantle convection to involve vertical motion alone . There must also be horizontal flow, which would cause plumes to tilt so that their surface expression would generally lie some distance away from the vertical projections of their deepest points (Steinberger and O'Connell, 2000). This fact must complicate attempts to associate surface phenomena with deep structure accurately.

• Some of the reasoning about the geological significance of LLSVPs is circular. The quantitative correlation of LLSVPs and LIPs is based, among other things, upon an assumption of stability in the deep mantle over a period of 200 million years (Torsvik et al., 2006), but then (Burke et al., 2008) adduce evidence of large eruptions in the Archean as evidence for the persistence of deep-mantle PGZs for 2500 million years.

**Hindsight Heresy**

Everyone (well, almost) has noticed the occurrence of seemingly unlikely events:

• John Adams and Thomas Jefferson, second and third presidents of the United States, died on the same day, July 4, 1826, which was also the 50th anniversary of the Declaration of Independence.

• Abraham Lincoln's secretary was named Kennedy. John F. Kennedy's secretary was named Lincoln.

Few people think that such coincidences are evidence of paranormal phenomena, however (we hope). Among the huge number of events that occur in life, it is actually likely that a few will be highly unlikely. Acton (1959, Chapter 6) has given the name "Hindsight Heresy" to the logical error of assessing the improbability of an event using the same data sample that caused us to notice the event in the first place. Such reasoning could lead us to assign probabilities of less than $10^{-8}$ to the Adams-Jefferson-Independence coincidence, and about $10^{-7}$ to the Lincoln-Kennedy coincidence.

Acton (1959) identified two dangers associated with *a posteriori* hypotheses:

1. "The analyst will usually be drawn to test extreme deviations of the data but will use techniques that do not presume any such selection."
2. "The analyst will mentally apply and reject *many* hypotheses before settling on the ones to be tested but will not adjust his error rate to correspond to his behavior."

*Extreme deviations* – To perform a meaningful *a posteriori* test one must estimate how many possible alternatives there were to the outcome that was observed, and this is difficult to do objectively. In the Lincoln-Kennedy case, for example, we might consider all the president-secretary pairs. But we probably should also consider other pairings that would have caught our attention, such as ones involving presidential family members, political allies or rivals, etc. And in addition to presidents, perhaps we should include movie stars, monarchs, or other famous people. Because of the nature of combinatorial functions, the number of possible pairings grows very rapidly with the number of individuals considered, so seemingly astronomical improbabilities may not actually be so.

*Unconscious Bias* – Even if one is able to use independent data to assess a hypothesis, there is still a great risk of self-delusion. There is typically a choice of data sets, and of analysis methods to use, and of possible changes to the hypothesis under test, which can lead to deceptively good "agreement" between theory and data.

"Hindsight Heresy" errors are a particular danger in Earth science, where it is seldom possible to conduct controlled experiments.

As seismologists, we don't have to look far for an example: In 1985, the U. S. Geological Survey made a formal prediction that, with a confidence level of 95%, an earthquake of magnitude about 6 would occur on the San Andreas fault near Parkfield, California, before 1993 (Bakun and Lindh, 1985). It based this prediction upon a sequence of "repeating" Parkfield earthquakes that had occurred in 1857, 1881, 1901, 1922, 1934, and 1965. Inasmuch as such regularity is rare (nearly unique, in fact) in earthquake catalogs, this prediction would seem to be a likely example of a hindsight error, and indeed the predicted earthquake did not occur. The next significant Parkfield earthquake, of magnitude 6.0, occurred on September 28, 2004, but in the wrong place, and of course much later than predicted.

Even more disturbingly, it turned out that the statistical confidence limits given for the time of the predicted earthquake had in effect been based upon a *modified* earthquake catalog, in which the 1934 event was moved to 1943, because of a supposed perturbing effect of a foreshock (Savage, 1993). This modification of the catalog resulted in an even more regular pattern of repeating events, and thus to the high confidence claimed for the time of the predicted event. Few seismologists became aware of this chain of reasoning until the earthquake was already overdue.

## LARGE LOW-SHEAR-VELOCITY PROVINCES (LLSVPs)

Among the first features of Earth's deep mantle to be reliably resolved by seismic tomography was a large volume having low shear-wave speed ($V_S$) located near the bottom of the mantle beneath southern Africa. Subsequently, another low-$V_S$ region was found at an approximately antipodal position beneath the southern Pacific Ocean. These features have horizontal dimensions as great as 135° (7500 km at the core-mantle boundary) and extend upward from the boundary about 500 km. For a while, they were "commonly referred to as superplumes" (Romanowicz and Gung, 2002), but accumulating observations have shown that this name is inappropriate and misleading.

### Elastic Moduli

The anomalies in the compressional-wave speed, $V_P$, in the African and Pacific LLSVPs are much weaker than the $V_S$ anomalies. In terms of elastic moduli, these seismic wave speeds are

$$V_P = \sqrt{\frac{k + (4/3)\mu}{\rho}} \tag{1}$$

and

$$V_S = \sqrt{\frac{\mu}{\rho}} \ , \tag{2}$$

where $k$ is the bulk modulus (incompressibility), $\mu$ is the rigidity modulus (shear stiffness), and $\rho$ is density. From the seismic wave speeds it is possible to compute the "bulk sound speed",

$$V_\phi = \sqrt{V_P^2 - (4/3)V_S^2} = \sqrt{\frac{k}{\rho}} \ , \tag{3}$$

which does not correspond to any possible elastic wave in a solid, but reflects the medium's resistance to isotropic volume changes involving no shear deformation. It turns out that the LLSVPs have *positive* $V_\phi$ anomalies (Figure 2, right). Because all three types of wave speed have negative derivatives with respect to temperature $T$,

$$\frac{\partial V}{\partial T} < 0 \ , \tag{4}$$

the LLSVPs cannot be caused by high temperature alone, must involve compositional or phase heterogeneity, and might not involve temperature at all.

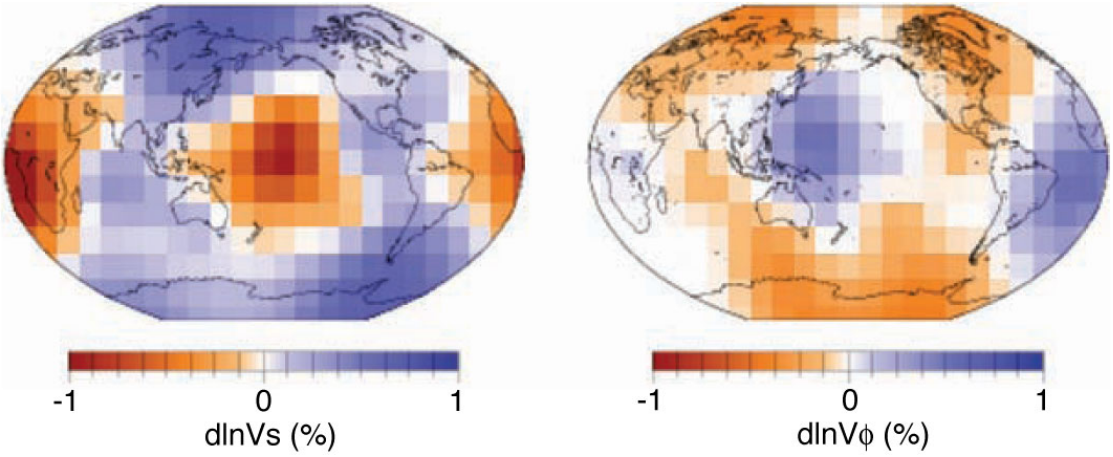


Figure 2: Maps of horizontal variations of the shear-wave speed, $\Delta V_S$, (left) and of the bulk sound speed, $\Delta V_\phi$, (right) in the deep mantle (depth 2000 to 2891 km), from Trampert *et al*. (2004). The opposite signs of the anomalies indicate that the LLSVPs are not simple temperature effects, but must involve compositional or phase differences.

**Density**

At high frequencies, where ray theory is a good approximation, seismograms depend only upon the distribution of seismic-wave speeds in the Earth (and the source locations and characteristics, of course). More generally, however, seismic waves are also sensitive to density within the Earth for wavelengths comparable to or larger than the sizes of heterogeneities. A familiar example of this fact is given by the reflection and transmission coefficients of elastic waves at sharp interfaces (*e.g.* Aki & Richards (1980), Sections 5.2.3, 5.2.4) which depend on both the densities and the wave speeds. Earth's normal modes, which have frequencies as low as about one cycle per hour, are sensitive to the density distribution and provide another example. Observations of the splitting of overtone modes (Trampert et al., 2004) show that the LLSVPs are characterized by *high* densities relative to their surroundings (Figure 3). They are thus not buoyant, and the name *superplume* is misleading.

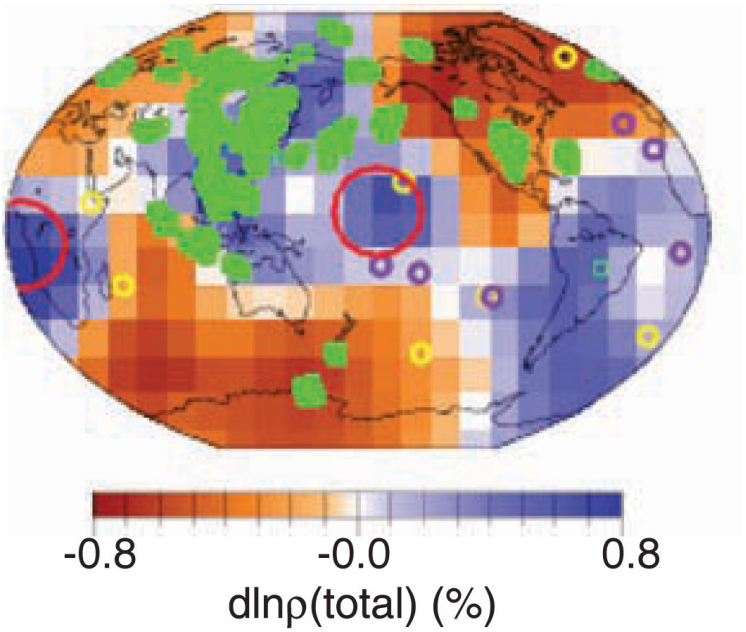


Figure 3: Relative density variations in the deep mantle (depth 2000 to 2891 km), from Trampert et al. (2004). Red circles indicate the locations of the two LLSVPs, which are regions of *high* density and thus are gravitationally stable.

**Possible Errors**

Are the seismic-wave speed models, upon which these conclusions are based, reliable? Could either random errors, caused by data inadequacy, or systematic errors, caused by neglect of relevant physical effects, be large enough to produce the apparent anti-correlation of $\Delta V_S$ and $\Delta V_\phi$ in the deep mantle?

The sampling of the mantle by seismic waves is far from complete, because of the uneven geographical distribution of earthquakes and seismometers, so derived three-dimensional models must be inaccurate and contain artifacts. Della Mora et al. (2011) estimated the magnitude of these errors by inverting synthetic arrival-time data whose distribution mimics that of real data, and concluded that the effects of data distribution and noise are not adequate to explain the observations.

Might the seismic wave speed models be subject to some kind of systematic error? Anelasticity can slightly but significantly affect wave speeds, and might conceivably increase the sensitivity of $V_S$ and $V_\phi$ to temperature, so that the anomalies shown in Figure 2 could be temperature effects after all. Current knowledge of the properties of perovskite under deep-mantle conditions indicates that anelastic effects are probably not strong enough to change the conclusion that the LLSVPs are chemically different from their surroundings (Brodholt et al., 2007).

But the detailed structure of the lower mantle remains an open question. Simmons et al. (2010) derived a three-dimensional mantle model from seismic body-wave times, geodynamic observations, and mineral physics data. No low-frequency seismic data that would be sensitive to density were used. In this model, the LLSVPs are hotter than their surroundings, and both temperature and composition significantly affect $V_S$. The model

was "constructed with the underlying hypothesis that temperature variations are the dominant cause of mantle heterogeneity…", however, so it presumably has the strongest temperature anomalies that are consistent with observations. The persistence of compositional components of the LLSVPs in this model confirms that these features are chemically different from their surroundings, but the model also indicates that a contribution from temperature effects cannot be ruled out entirely.

## HOT SPOTS AND LIPS

### Inadvertent Selection

The most problematic of the "hindsight" effects identified by Acton (1959) is the inadvertent introduction of bias when formulating hypotheses. Introducing bias in this way is nearly impossible to avoid, and easy to overlook. For example, in hypothesizing that hot spots and LIPs occur above the margins of LLSVPs, and in defining a margins as the location of the $-1\%\,\Delta V_S$ contour, one is excluding the more obvious hypotheses that they occur above the centers of LLSVPs (because that hypothesis doesn't fit as well) or outside of LLSVPs (for the same reason), and also excluding possible association with compressional-wave speed, bulk sound speed, density, or other physical properties, and with other contour levels.

Burke et al. (2004) used 25 LIPs to compute correlations with LLSVPs, and the later paper of Torsvik et al. (2006) used 23, both lists attributed to Coffin and Eldholm (1992). But the Coffin and Eldholm list contains 60 LIPs, and there are about $5\times10^{16}$ ways in which a 25-member subset of them could have been chosen.

- Which LIPs/hot spots to include? Many catalogs, of widely differing sizes, exist. A new one was generated.
- Which tomographic model? Three models were averaged, with differing weights.
- The definition of "margin" was a variable.
- Geometry: within anomalies?, outside anomalies?, near anomaly margins?, …
- Excluding worst-fit LIPs (*e.g.* Siberian Traps) from statistics(!).

### How Many Hot Spots Are There? And Where?

Even the identification of current-day hot spots is problematical. There are many published hot-spot catalogs, of widely differing sizes. In formulating their hypothesis, Torsvik et al. (2006) and Burke et al. (2008) apparently used subsets of their own choosing.

To minimize this problem, we use in our analysis five different hot-spot catalogs, taken from (Courtillot et al., 2003), (Morgan and Morgan, 2007), (Richards et al., 1988), (Sleep, 1990), and (Steinberger, 2000), containing from 37 to 72 hot spots each and with the locations of particular hot spots sometimes differing slightly from catalog to catalog (Figures 4a to 4e).

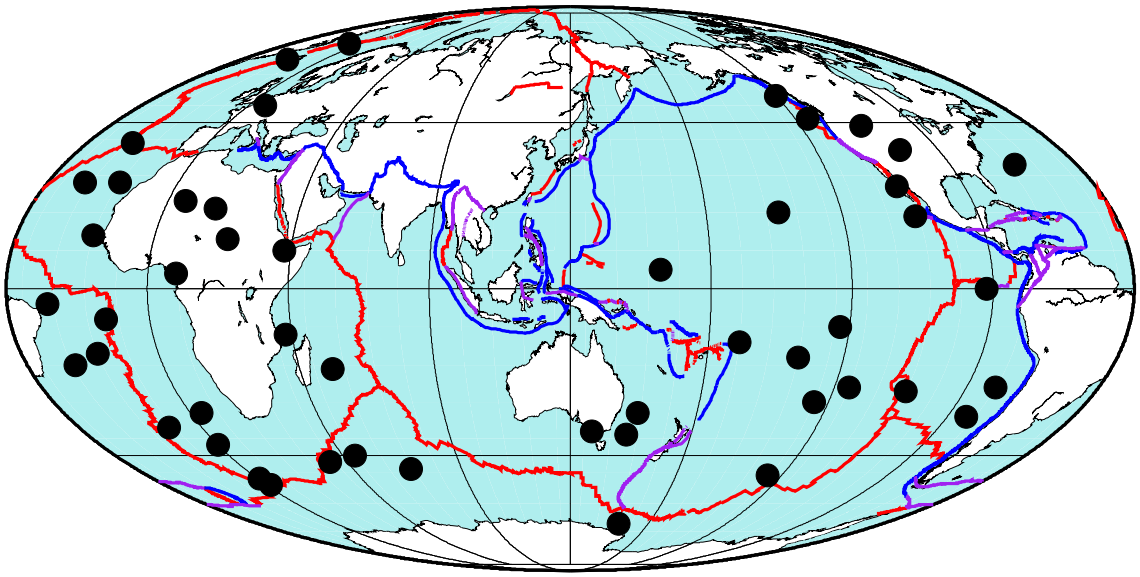

Figure 4a: Map of the Earth (Hammer equal-area projection), showing the 49 hot spots "found in the most cited catalogues", from Table 1 of Courtillot *et al*. (2003). Red, purple, and blue lines: Spreading, strike-slip, and convergent plate boundaries (Coffin et al., 1998).

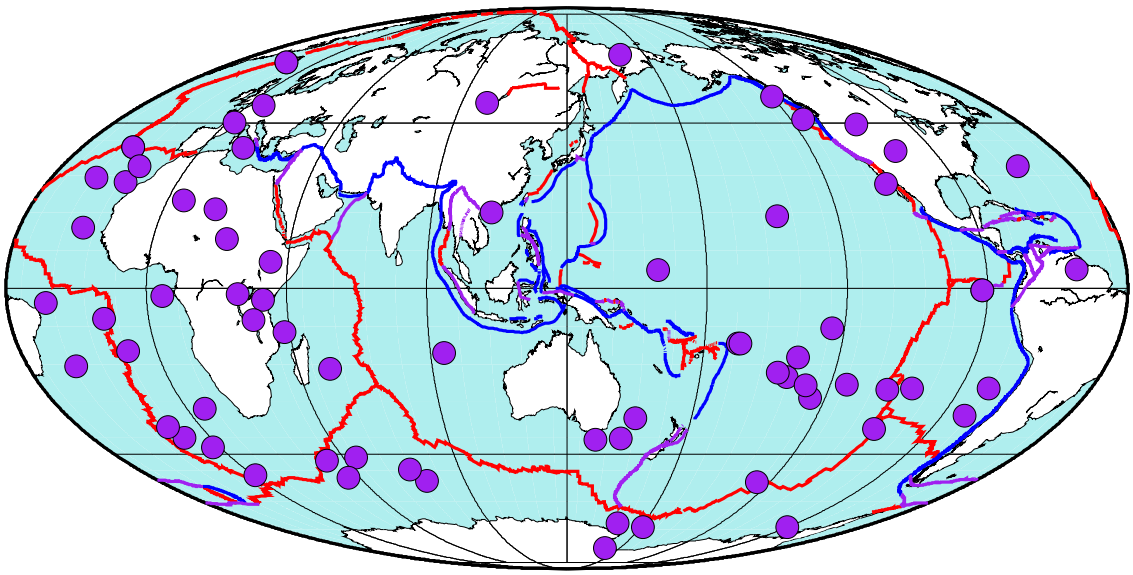

Figure 4b: The 72 hot spots given in Table 1 of Morgan and Morgan (2007).

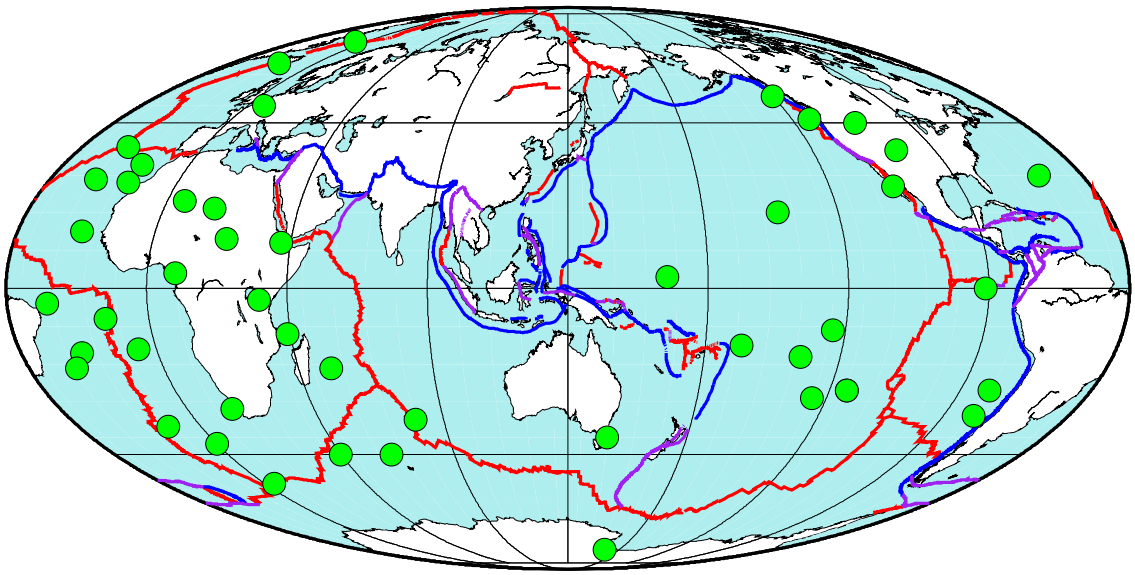

Figure 4c: The 47 hot spots given in Table 1 of Richards et al. (1988).

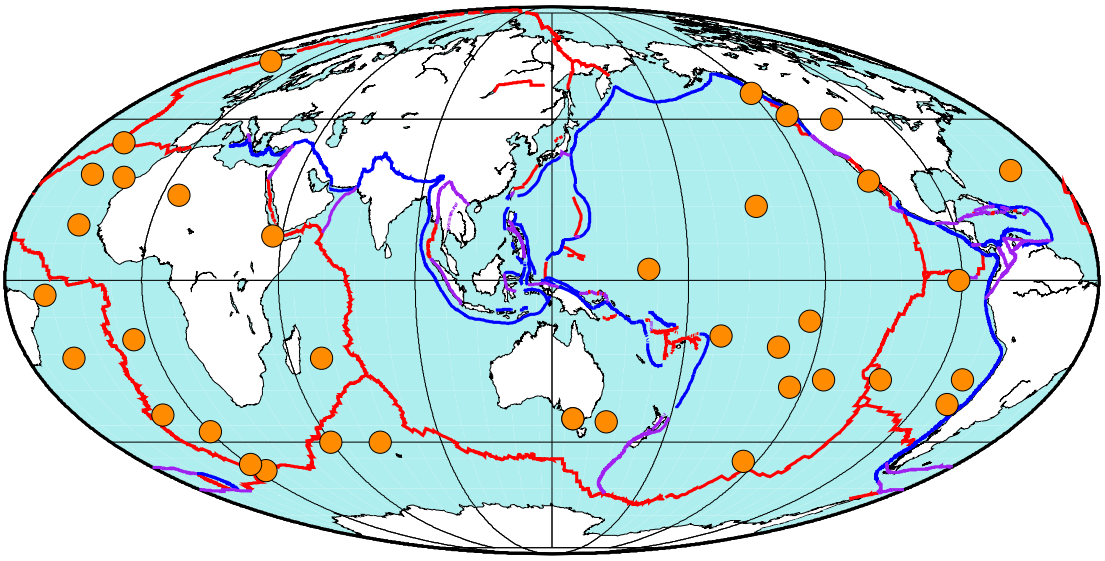

Figure 4d: The 37 hot spots given in Table 1 of Sleep (1990). Hot-spot coordinates taken from Richards et al. (1988) or from Courtillot et al. (2003), on the advice of Sleep (personal communication, 2014).

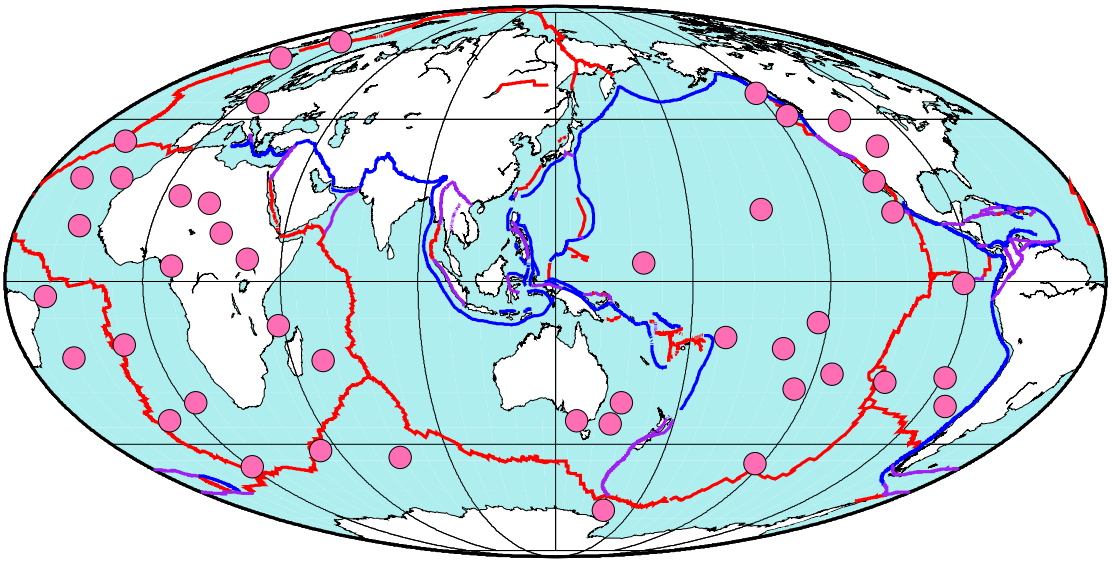

Figure 4e: The 44 hot spots given in Table 1 of Steinberger (2000).

Many apparent departures from randomness are evident on these maps. For example, there are few hot spots in Asia or the eastern Indian Ocean. Sleep's list has none between longitudes of 65° and 143°, a lune that includes about 21.7% of Earth's surface. The probability that 37 uniformly distributed random points would all lie outside such an area is $(1-0.217)^{37} = 1.2\times10^{-4}$. In the lists of Courtillot et al. (2003), Richards et al. (1988), and Steinberger (2000), the corresponding lunes are slightly smaller but the total numbers of hot spots are greater, so the probabilities are $1.3\times10^{-5}$, $3.9\times10^{-5}$, and $4.0\times10^{-5}$. Of course, the main reason that these probabilities are small is that we chose the longitude boundaries to make them small. Much larger departures from randomness could be found if we designed regions with boundaries more complicated than meridians. Clearly, catalogued hot spots are not distributed globally in a uniform random manner. But do the departures from randomness reflect limitations in our geological knowledge (*e.g.* of Asia), systematic errors of analysis, or processes in the deep mantle?

## CORRELATION WITH SPREADING PLATE BOUNDARIES

The maps in Figure 4 also suggest that hot spots occur more frequently near boundaries where tectonic plates are spreading than elsewhere. This example is well suited to illustrating assessment of correlations involving greater geometrical complexity than the meridian-bounded case analyzed above.

To assess confidence in a correlation between the locations of features such as hot spots and a curve such as a tectonic plate boundary, we first need to know the statistics of the distances from random locations to the curve. Figure 5 shows examples of such information, in the form of statistical cumulative distribution functions (CDFs), of the distance from random locations to each of the Earth's three types of tectonic plate boundaries. These functions depend on the lengths and irregular shapes of the plate

boundaries, and were obtained empirically using a Monte Carlo method, in which we generated many (10,000) pseudo-random locations, statistically uniformly distributed over the globe, and for each location computed the distance to the nearest point on each type of plate boundary. We used the digital plate boundaries of Coffin et al. (1998), shown in Figures 4a-e, which consist of many separate segments defined by discrete points. These contain 188 separate spreading segments, for example, with a total length of 112,897 km and a mean spacing between sample points of about 19 km. The convergent and transform boundaries are 75,464 km and 40,280 km long, but their shapes are such that their CDFs happen to be rather similar to each other. The curves in Figure 5 show, for example, that almost 50% of Earth's surface is within 10° of a ridge, and that all of Earth's surface is within about 38° of a ridge.

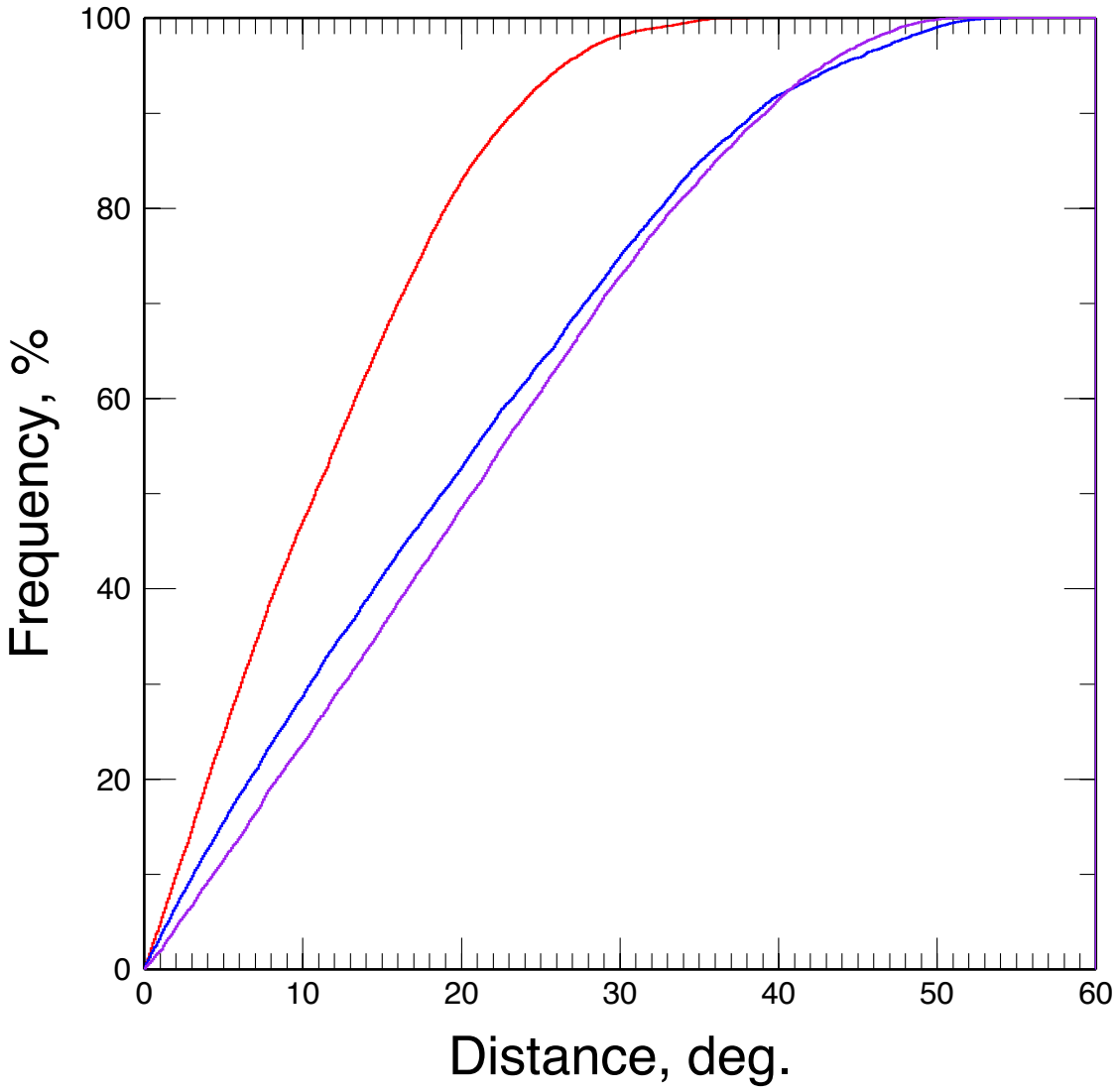


Figure 5: Empirical cumulative distribution functions (CDFs) for the distance from a random location on the globe to tectonic plate boundaries of different types. Red: spreading boundaries ("Ridges"); Blue: convergent boundaries ("Trenches"); Purple: strike-slip boundaries ("Transforms"). Each curve is a cumulative histogram of the values of the central angles between

each of 10,000 pseudo-random locations, uniformly distributed statistically over the Earth, and the nearest point on a plate boundary of the appropriate type. Digital plate boundaries taken from Coffin et al. (1998).

Curves such as those in Figure 5 give the probability $p$ that a random location is within a specified distance of a particular curve on Earth's surface. The probability of $k$ random locations being within this distance, out of $n$ locations in a catalog, then follows the binomial distribution:

$$b(k;n,p) = \binom{n}{k} p^k (1-p)^{n-k} \quad , \tag{5}$$

where

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} \tag{6}$$

is the binomial coefficient (the number of $k$-member subsets of $n$ distinct objects). We are interested, however, not in the probability of *exactly* some number $m$ of "successes", but in the probability of *at least* $m$ successes, $1 - B(m-1;n,p)$, where

$$B(m;n,p) = \sum_{k=0}^{m} b(k;n,p) \tag{7}$$

is the cumulative binomial distribution function.

Figure 6 shows the results of applying this type of analysis to the distances from hot spots to spreading plate boundaries. The color-coded step functions on the left-hand side of the figure are the cumulative distributions of ridge-distances for the five hot-spot lists. From about 5° to 30°, the distributions for most of the catalogs lie above the "whole Earth" CDF for random locations. For example, 27 of Courtillot's 49 hot spots (55%) lie within 7.21° of a ridge, whereas the expected fraction for random locations is about 34.9%.

This analysis seems to confirm that hot spots tend to be closer to ridges than expected for random locations. We have based this assessment, however, upon the one particular point out of 49 on the distribution of Courtillot's hot spots that gives the lowest probability and thus the highest confidence. By choosing different points, we can obtain probabilities between 0.12% and 17% (confidence levels from 99.9% to 83%), and others of the five hot-spot catalogs give probabilities ranging from 0.012% to 75% (confidence levels from 99.99% to 25%). The curves on the right-hand side of Figure 6 illustrate this variation quantitatively. Each curve shows, for one hot-spot list, the cumulative probability $1 - B(m-1;n,p)$ of at least as many random locations lying as close to a ridge as actually occurs for each hot spot on the list. As we would expect, points that are farther above the "whole Earth" curve on the left correspond to smaller probabilities given by the curves on the right. In the example just mentioned, $n = 49$, $p = 0.349$, and the probability of at least 27 successes is about 0.12%.

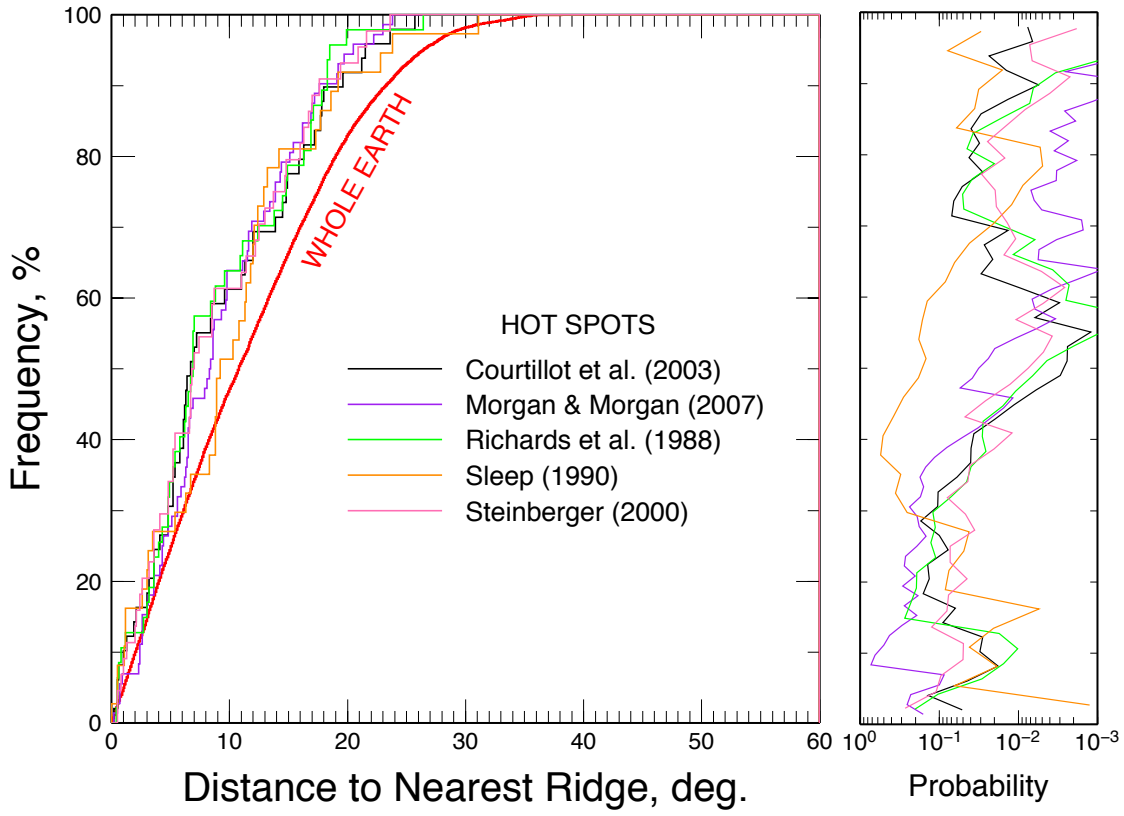


Figure 6: Left: Cumulative distributions of the spherical distances from hotspots to spreading plate boundaries ("Ridges"). Stair-step curves: distributions for five published hot-spot catalogs, color-coded as in the maps of Figure 4; heavy red curve: Cumulative distribution function for a random point on Earth's surface (Figure 5). Hot spots tend to be closer to ridges than expected at random. Right: For each hot-spot list, the probability of at least as many randomly chosen points being as close to the nearest ridge as each hot spot on the list. Values further to the right indicate higher confidence in the observed correlation.

Clearly, choosing the most favorable point from each hot-spot catalog gives a misleading idea of the true confidence level. To correct for this effect, we must multiply the probability by the number of alternatives from which it was chosen, that is, by the number of hot spots in the catalog. The right-hand column of Table I gives the probabilities that result from this correction, which correspond to confidence levels from 92.1% to 99.4%. Hot spots in the table of Richards et al. (1988) show the strongest correlation, but again, if we choose the most favorable catalog out of five, we must multiply the probability by five. Thus we conclude that hot-spot locations are correlated with spreading ridges at a confidence level in the range of about 92% to 97%.

**TABLE I – Probabilities of Hot-Spot Spreading-Ridge Correlation**

| Catalog | No. of Hot Spots | p (extreme) | p (normalized) |
|---|---|---|---|
| (Courtillot et al., 2003) | 49 | 0.0012 | 0.059 |
| (Morgan and Morgan, 2007) | 72 | 0.00036 | 0.026 |
| (Richards et al., 1988) | 47 | 0.00012 | 0.006 |
| (Sleep, 1990) | 37 | 0.0013 | 0.048 |
| (Steinberger, 2000) | 44 | 0.0018 | 0.079 |

## CORRELATION WITH DEEP-MANTLE STRUCTURE

We use the same type of analysis that was used above for spreading plate boundaries to assess the correlation between hot spots and the margins of the LLSVPs in the deep mantle.

Figures 7 and 8 show maps of the variation of the shear-wave speed $V_S$ at a depth of 2800 km (roughly 90 km above the core boundary) in two recent three-dimensional mantle models, of Kustowski et al. (2008) and Ritsema et al. (2011). Both models are based on tens of thousands of observed body-wave arrival times, surface-wave speeds, and normal-mode frequencies and splitting functions, and the differences between the models represent uncertainties caused by the incomplete data distributions and different analysis methods. The dominant features on both figures are the two LLSVPs centered beneath southwest Africa and the central Pacific Ocean. These negative $\Delta V_S$ anomalies are similar, though not identical, in shape in the two models, but differ significantly in their strengths.

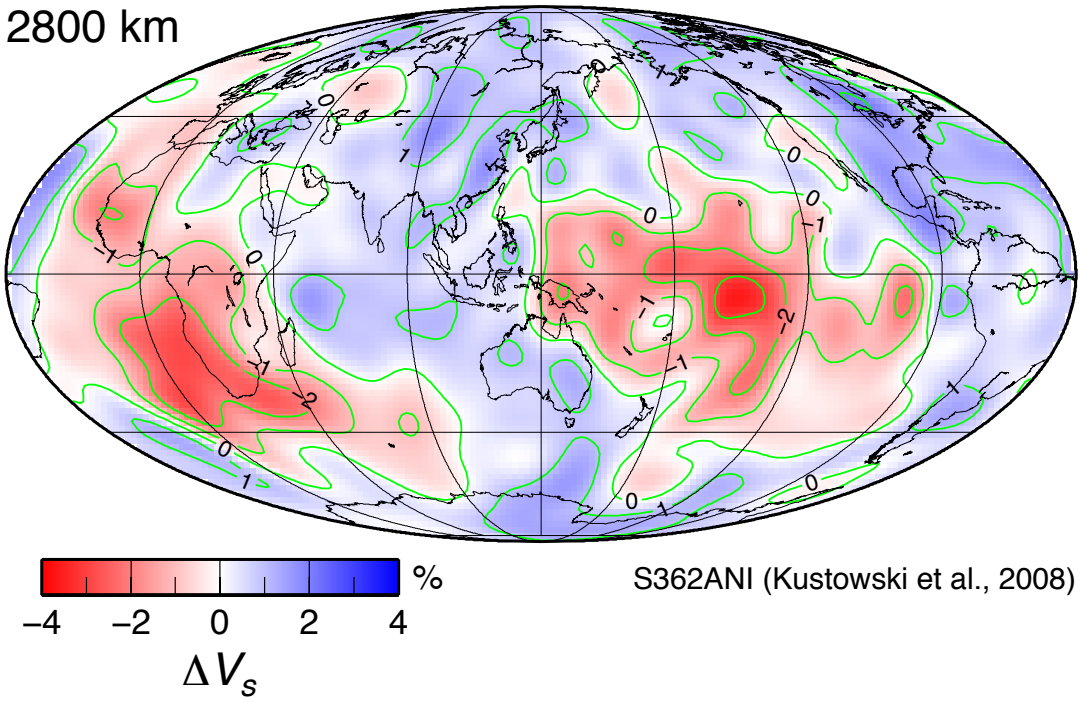


Figure 7: Geographical variation in shear-wave speed near the base of the mantle (depth 2800 km) in the three-dimensional tomographic model S362ANI of Kustowski et al. (2008). This model is anisotropic; the value plotted corresponds to the Voigt average of the shear moduli. Green curves: contours (interval 1%).

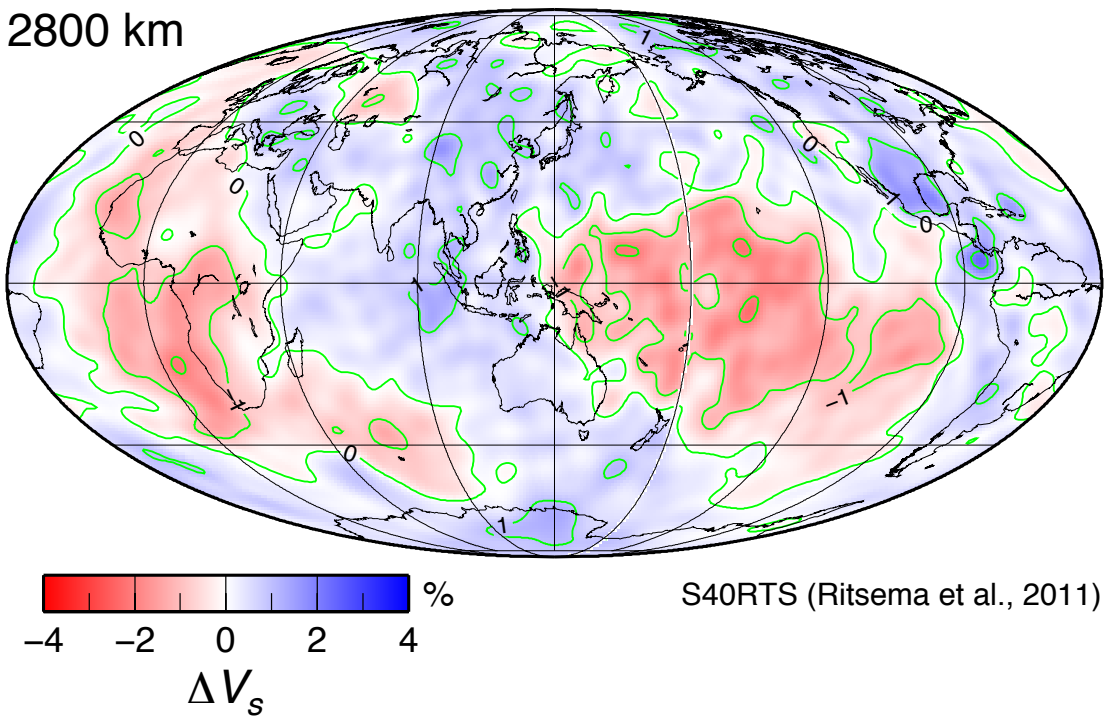


Figure 8: Like Figure 7, for $V_S$ variations in the model S40RTS of Ritsema et al. (2011)

Figures 9 and 10 show the cumulative distribution functions of angular distance from random points on Earth's surface to the −1% $\Delta V_S$ contour in each of these mantle models. These are analogous to the distribution functions for plate boundaries shown in Figures 5 and 6, and were obtained using the same Monte Carlo method. About 50% of Earth's surface is within 18.1° (S362ANI) or 18.5° (S40RTS) of the −1% $\Delta V_S$ contours. The color-coded step functions in Figures 9 and 10 show the empirical distance distributions for hot spots in each of the five catalogs we consider, analogous to the distributions for spreading-ridge distances in Figure 6. Also like Figure 6, the right-hand side plots on these figures show, for each point on the empirical distribution for each hot-spot catalog, the probability of as many or more randomly chosen points being as close to the contour as hot spots in the catalog are.

For all mantle models, hot-spot catalogs, and distances, hot spots tend to lie closer to the −1% $\Delta V_S$ contours than expected at random. Tables I and II give the probabilities associated with the most extreme points on the probability plots. As above, the "normalized" probabilities, which account for the number of hot spots in each catalog, give the most meaningful measures of confidence. The probabilities of the observed correlations lie in the approximate range 0.001 to 0.004 (confidence levels 99.9% to 99.4%), with model S362ANI giving somewhat higher confidence levels. These probabilities are 2½ to 6 orders of magnitude higher than those given by Burke et al. (2008, Fig. 5), due, probably, to the efforts we have taken to avoid or account for biasing effects.

The Burke et al. probabilities are nevertheless low enough to reject the null hypothesis (that the correlation is just a coincidence) at a confidence level exceeding 99%. But correlation is not causation. Statistical tests of this sort can give us confidence that a null hypothesis is incorrect, but do not tell us what hypothesis may be correct.

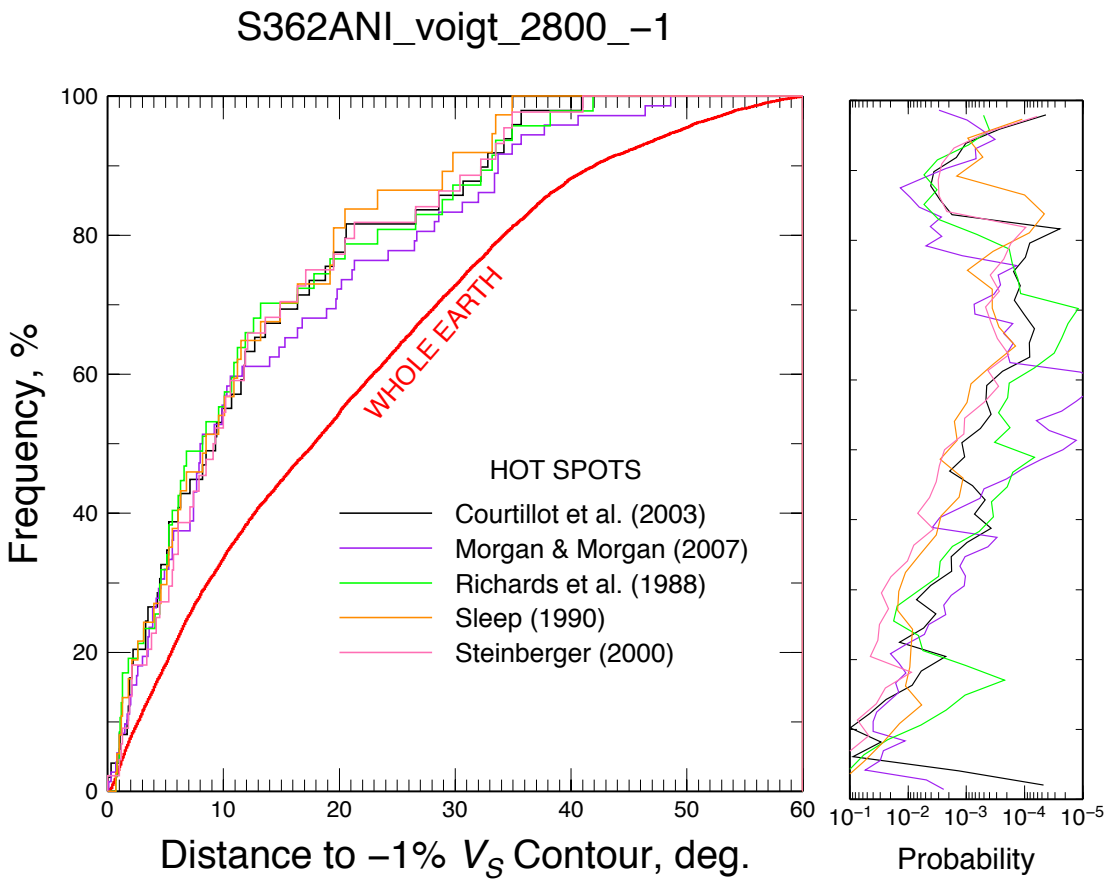


Figure 9: Left: Distribution functions for the spherical distances from hotspots to the −1% $V_S$ contour at 2800 km depth in mantle model S362ANI of Kustowski et al. (2008). Stair-step curves: the distributions for five published hot-spot lists, color coded as in Figures 4 and 6; heavy red curve: the distribution function for a random point on the globe. Hot spots tend to be closer to the −1% $V_S$ contour than expected at random. Right: For each hot-spot list, the probability of as many or more randomly chosen points being as close to the nearest contour as are points on the list. Values further to the right indicate higher confidence in the observed correlation.

**TABLE II – Probabilities of Observed Hot-Spot $V_S$ Correlation: Model S362ANI**

| Catalog | No. of Hot Spots | p (extreme) | p (normalized) |
|---|---|---|---|
| (Courtillot et al., 2003) | 49 | 0.000025 | 0.0012 |
| (Morgan and Morgan, 2007) | 72 | 0.000004 | 0.0003 |
| (Richards et al., 1988) | 47 | 0.000012 | 0.0006 |
| (Sleep, 1990) | 37 | 0.000046 | 0.0017 |
| (Steinberger, 2000) | 44 | 0.000061 | 0.0026 |

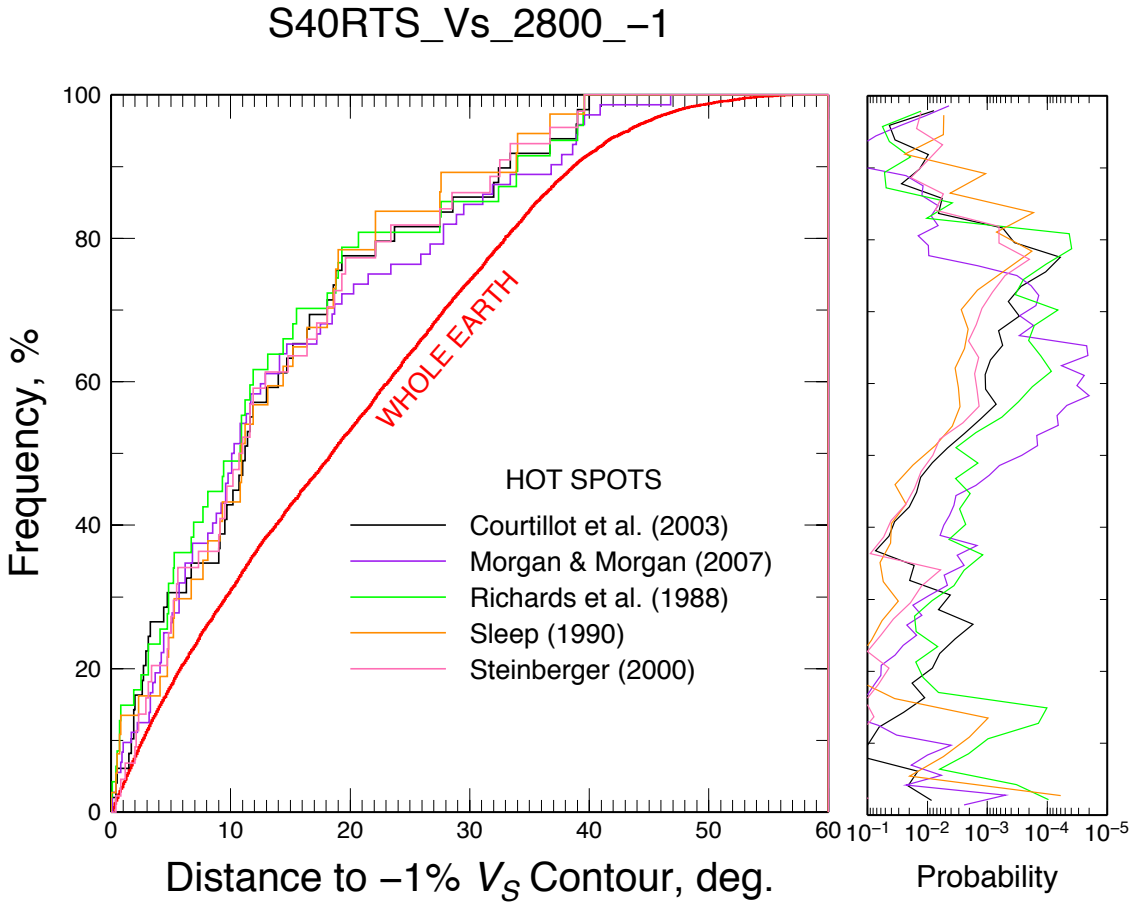


Figure 10: Like Figure 9, for mantle model S40RTS of Ritsema et al. (2011).

**TABLE III – Probabilities of Observed Hot-Spot $V_S$ Correlation: Model S40RTS**

| Catalog | No. of Hot Spots | p (extreme) | p (normalized) |
|---|---|---|---|
| (Courtillot et al., 2003) | 49 | 0.000061 | 0.0030 |
| (Morgan and Morgan, 2007) | 72 | 0.000020 | 0.0014 |
| (Richards et al., 1988) | 47 | 0.000039 | 0.0018 |
| (Sleep, 1990) | 37 | 0.000169 | 0.0063 |
| (Steinberger, 2000) | 44 | 0.000199 | 0.0088 |

Thus it appears that significant correlations can be found between hot-spot locations and (1) simple regions bounded by meridians, (2) spreading plate boundaries, and (3) certain contours of shear-wave-speed anomalies near the bottom of the mantle. As these examples illustrate, however, this apparent significance can be produced, or greatly amplified, by inadvertent "hindsight" effects.

## KIMBERLITES AND ASSOCIATED VOLCANISM

In a more recent paper, Torsvik et al. (2010) argue that kimberlite eruptions, and thus diamond deposits, occur preferentially within 15° of the edge of the African LLSVP, defined as the −1% $\Delta V_S$ contour in the deep mantle. They claim that about 80% (1112 out of 1395) kimberlites and related volcanics erupted since about 320 Ma lie within this zone, and therefore propose that, in addition to causing hot spots and LIPs, deep-mantle plumes cause this distinctive style of volcanic activity, comprising clusters of pipes, dykes and sills. These 15° zones cover 42% to 45% of Earth's surface, however, raising the same questions we discussed above regarding the likelihood that the correlations occur by chance. These workers' extension of their hypothesis to a different type of volcanism is salutary, because it makes possible tests using independent data. Nevertheless, even if such a broad definition of the edge of an LLSVP is accepted, there are serious geological objections to the proposal.

As Torsvik et al. (2010) note, the Late Cretaceous/Early Tertiary kimberlites of the Slave Province, Canada, which represent 17% of their database, do not fall within the 15° zones. Another exception, West Kimberley Province of NW Australia, contains over 100 Miocene (about 18-25 Ma) lamproites and associated volcanics, including numerous diamond-bearing pipes (Jaques et al., 1986). Figure 11 shows the location of this province relative to the 30-Ma time slice of Torsvik et al. (2010). Pipes emplaced in the north of Australia at about 240 Ma (Torsvik et al., 2010, Supplementary data) constitute another exception. Clearly, proximity to a LLSVP is not required to trigger kimberlite and related alkaline volcanism.

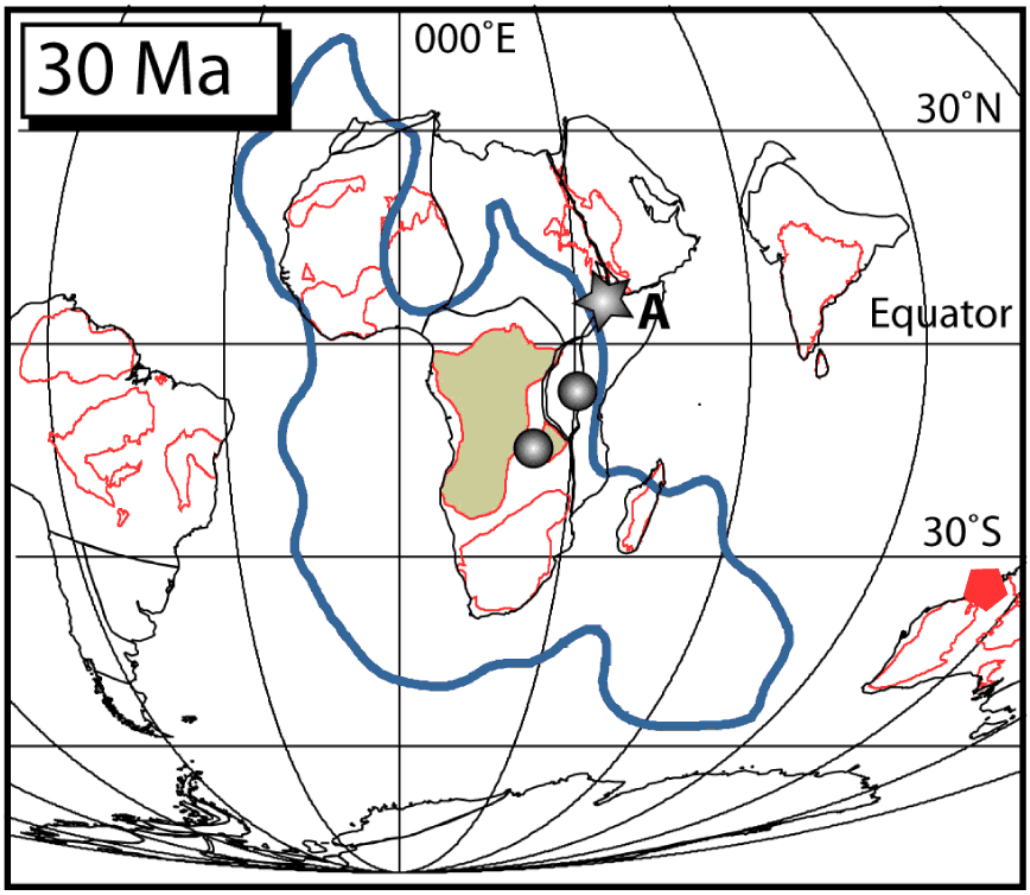


Figure 11: Distribution of kimberlite clusters (circles) relative to the African LLSVP (blue line) at 30 Ma, after Torsvik et al., (2010). Star (A): the Afar LIP; red pentagon: the location of the Miocene (~17-25Ma) West Kimberley lamproite province, NW Australia, with over 100 individual pipes, several of which are diamond bearing. This latter province is located well beyond the 15° zone surrounding the African LLSVP.

Of the kimberlites considered by Torsvik et al. (2010), about 80% erupted in the period 120-50 Ma, with more than 60% being younger than 100 Ma. The dataset is dominated by Africa, which hosts more kimberlite clusters than any other continent and experienced widespread, repeated kimberlite volcanism immediately before and after the disruption of Gondwana. Africa drifted slowly northeastward during this time, and so stayed close to the Africa LLSVP. A comprehensive database of the ages of kimberlites and related volcanics is available, particularly for southern Africa, providing excellent data to test the proposed correlation.

Figure 12 shows the age distribution of alkaline volcanics erupted in southern Africa during the period leading up to and following disruption of Gondwana (data from Moore et al. (2008)). Jelsma et al. (2009; 2004) give similar age distributions. The period preceding continental break-up was characterized by minor intermittent kimberlite volcanism, followed by a major increase in activity shortly after the start of continental rifting and the opening of the Indian and Atlantic Oceans. Several distinct episodes of kimberlitic and allied volcanism followed, each preceded by a reorganization of the

spreading regime of the ocean ridge system surrounding southern Africa. The correlation between episodes of post-Gondwana alkaline volcanism in southern Africa and spreading-ridge reorganizations suggests a tectonic trigger linked to long-range intra-plate transmission of compressive stresses (Moore et al., 2008, 2009). It is not clear how an LLSVP source might account for the episodic nature of post-Gondwana kimberlite activity, or for the correlations with spreading-ridge tectonics.

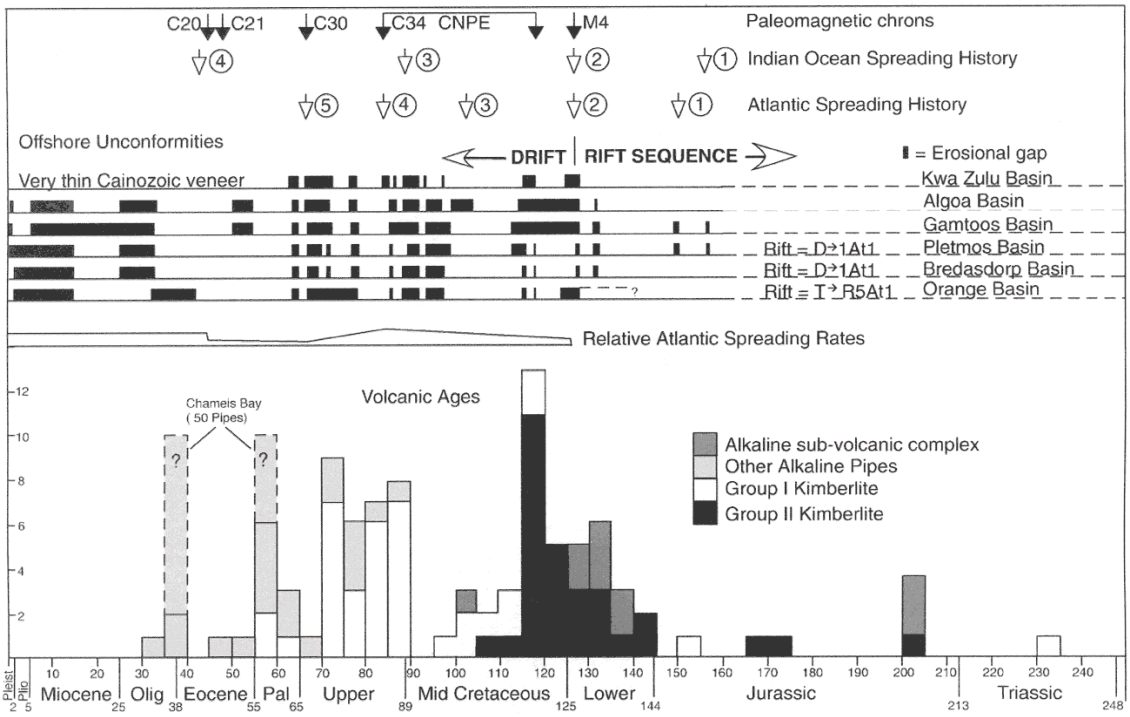


Figure 12: A summary of volcanic ages, offshore unconformities (McMillan, 2003) and Atlantic spreading histories (Nürenberg and Müller, 1991). CNPE = Cretaceous Normal Polarity Episode. Indian Spreading History from McMillan (2003) and Reeves and de Wit (2000): 1 – Initial rifting between Africa and Antarctica; 2 – Commencement of spreading; 3 and 4: Changes in Indian Spreading regime recognized by Reeves and de Wit (2000). Atlantic spreading history from Dingle and Scrutton, (1974) and Nürenberg and Müller (1991); I – Rifting extends into southern Atlantic Ocean; 2 – Commencement of opening of Atlantic (drift sequence)(M4); 3 – Estimated time of separation of Falkland Plateau and Agulhas bank, based on assumed spreading rates; 4 – Major shift in pole of rotation of Africa/South American plates; 5 – Beginning of progressive shift in pole of rotation of African/South American plates. Reeves and de Wit (2000) suggested that Atlantic spreading commenced earlier (~136 Ma) than the timing (~127 Ma) inferred by Nürenberg and Müller (1991) and McMillan (2003). Dashed lines and question marks for the Chameis Bay pipes denote the two different ages indicated by field relationships and very limited radiometric dating. Further work is required to establish the respective numbers of pipes of each age.

Most kimberlites erupted in southern Africa prior to about100 Ma were Group II types (also termed Orangeites), which are usually relatively low-volume dykes and small blows covering at most a few hectares. Subsequent to about 100 Ma, the kimberlites were Group I varieties, which are typically much larger. For example, the Orapa kimberlite

covers a surface area of about110 ha. Torsvik et al. (2010) show a series of 10 Ma time slices for the period 250-30 Ma to illustrate the relationship between the African LLSVP and the distribution of kimberlites and associated volcanics. During this period, much of the continent was close to the African LLSVP. It is thus inevitable that kimberlites erupted in Africa over this period lie near the edge of the African LLSVP.

The data presented by Torsvik et al. (2010) show a remarkably close association between kimberlite distribution and the edge (−1% $\Delta V_S$ contour) of the LLSVP during the period (200-150 Ma) leading up to the disruption of Gondwana. Figure 13a shows this association for 10-Ma time slices centered on 160 Ma and 150 Ma. This was, however, a period of very low kimberlite activity in Africa (Fig. 12). During the main episodes of kimberlite and allied volcanic activity that followed the break-up of Gondwana (about 120 to 30 Ma) the kimberlite distribution shows, in contrast, a wide geographical spread, ranging from near the LLSVP to far outside the 15° bands, often within the same 10-Ma period (Fig. 13b). The scatter distant from the LLSVP is particularly extreme for the time slices centered on 100-70 Ma, which represent more than half of the kimberlites used by Torsvik et al. (2010).

Torsvik et al. (2010) do not provide the kimberlite age data that underpin their model, making an independent quantitative evaluation impossible. Nevertheless, a qualitative evaluation of the data presented in their supplementary Figures S2-S5 leaves little doubt that within Africa, the major spatial correlation of kimberlites in the period 250-50 Ma is simply with the Kaapvaal craton. This underlines the fact that the Kaapvaal craton is the most productive in the world for both numbers of kimberlites and diamonds.

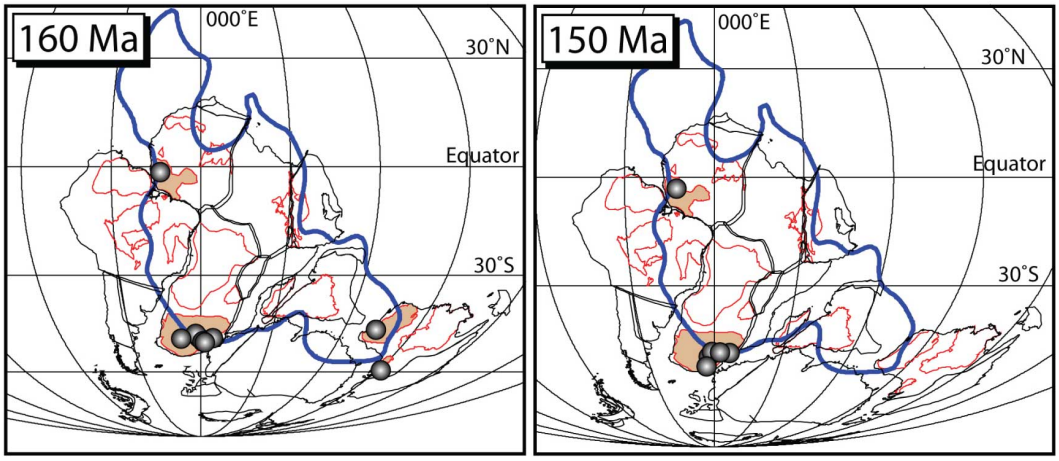


Figure 13a: Kimberlite activity relative to the African LLSVP prior to the disruption of Gondwana. Circles: kimberlite clusters; red lines: outlines of old cratons; brown shading: cratons with kimberlites in the relevant time interval.

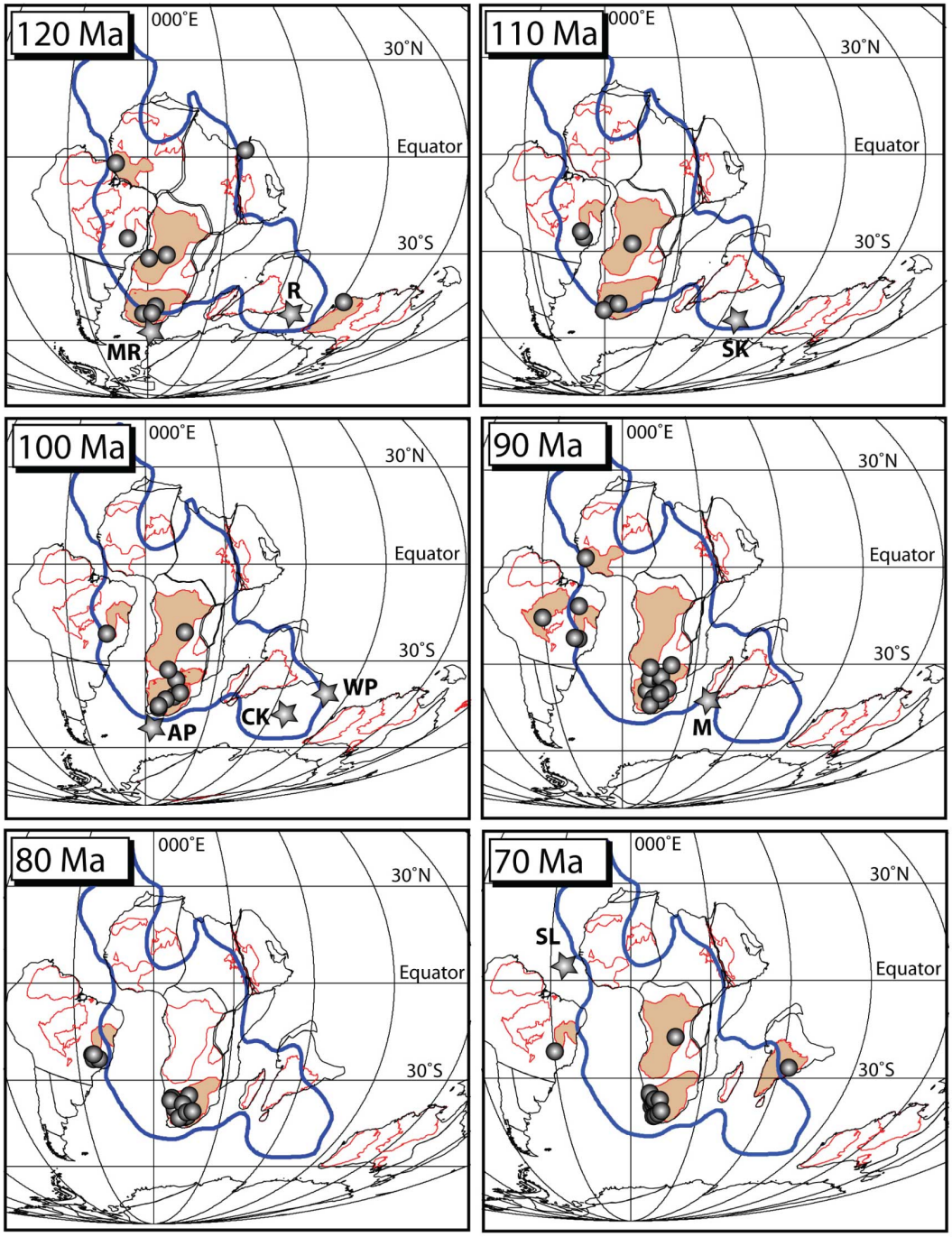


Figure 13b: Kimberlitic activity in Africa relative to the LLSVP subsequent to the disruption of Gondwana. Circles: kimberlite clusters; stars: LIPs; red lines: outlines of old cratons; brown shading: cratons with kimberlites in the relevant time interval; AP: Agulhas Plateau LIP (~100 Ma); MR: Maud Rise LIP (~125Ma); R: Rajmahal LIP (~118 Ma); SK: South Kerguelen (~114 Ma); CK: Central Kerguelen (~100 Ma); WP: Wallaby Plateau (~96 Ma); M: Madagascar-Marion (~87 Ma); SL: Sierra Leone (~73 Ma).

Torsvik et al. (2010) stress the high correlation between kimberlite occurrence and the edges of the LLSVP in the time slice centered on 160 Ma (Torsvik et al., 2010, Figs 2, S1

and S3). However, 160 Ma was a period of minor kimberlite activity both in southern Africa (Fig. 12) and worldwide. The correlation between kimberlites and the LLSVP margin is much worse subsequent to 100 Ma, when by far the greatest number and largest volumes of kimberlites erupted. At this time, northward motion of Africa had carried the Kaapvaal craton away from the southern edge of the LLSVP, and yet the craton remained a major focus of kimberlite activity.

The apparent correlation between kimberlite distribution and the edge of the LLSVP thus appears to be a statistical artifact resulting from the position of the Kaapvaal craton. These observations bring into serious question the claimed causative link between the deep mantle and kimberlite and related activity. The clear evidence that Late-Cretaceous/Early Tertiary kimberlite activity in the Slave province of North America and Miocene lamproites in NW Australia are unrelated to either the African or Pacific LLSVPs supports this conclusion.

## CONCLUSIONS

Testing the statistical confidence level of a correlation of current hot-spot locations, or the former locations of LIPs or kimberlites, with the seismic wave-speed structure of the deep mantle unavoidably uses the same data that suggested a possible correlation in the first place. This "Hindsight Heresy" effect (Acton, 1959) can introduce strong bias into statistical tests, because the hypotheses involved will have been tailored to the data, but the tests will assume statistical independence.

Because of such bias, the correlation of the margins of "large low shear velocity provinces" near the bottom of the mantle with hot spots and former locations of LIPs and kimberlites is much less significant than the estimates of Burke et al. (2008).

Nevertheless, some of these correlations have confidence levels in excess of 99%. But correlation is not causation. The correlation gives us high confidence that the null hypothesis (it's just a coincidence) is false, but does not tell us what hypothesis is correct. Other comparable correlations exist with longitude ranges, the African tectonic plate, and spreading plate boundaries.

The geographic distribution of LIPs (Figure 1) suggests that the correlation is stronger with the African LLSVP than with the Pacific LLSVP, and also that the correlations are stronger on the eastern sides of both LLSVPs than on their western sides.

All these facts suggest that these correlations are symptoms of as-yet-unidentified processes.

*Acknowledgement* – We thank B. Steinberger for helpful comments.

## REFERENCES

Acton, F. S., 1959, Analysis of Straight-Line Data, New York, John Wiley & Sons, Inc., 267 p.:

Aki, K., and Richards, P. G., 1980, Quantitative Seismology. Theory and Methods, San Francisco, W. H. Freeman and Company, 932 p.:

Bakun, W. H., and Lindh, A. G., 1985, The Parkfield, California, earthquake prediction experiment: Science, v. 229, no. 4714, p. 619-624.

Becker, T. W., and Boschi, L., 2002, A comparison of tomographic and geodynamic mantle models: Geochem. Geophys. Geosyst., v. 3, no. 1, p. 1003.

Brodholt, J. P., Helffrich, G., and Trampert, J., 2007, Chemical versus thermal heterogeneity in the lower mantle: The most likely role of anelasticity: Earth Planet. Sci. Lett., v. 262, p. 429-437.

Burke, K., Steinberger, B., Torsvik, T. H., and Smethurst, M. A., 2008, Plume generation zones at the margins of large low shear velocity provinces on the core-mantle boundary: Earth Planet. Sci. Lett., v. 265, no. 1, p. 49-60.

Burke, K., and Torsvik, T. H., 2004, Derivation of large igneous provinces of the past 200 million years from long-term heterogeneities in the deep mantle: Earth Planet. Sci. Lett., v. 227, p. 531-538.

Castle, J. C., Creager, K. C., Winchester, J. P., and van der Hilst, R. D., 2000, Shear wave speeds at the base of the mantle: J. Geophys. Res., v. 105, no. B9, p. 21543-21557.

Coffin, M. F., and Eldholm, O., 1992, Volcanism and continental break-up: a global compilation of large igneous provinces, *in* Storey, B. C., Alabaster, T., and Pankhurst, R. J., eds., Magmatism and the Causes of Continental Break-up, p. 17-30.

Coffin, M. F., Gahagan, L. M., and Lawver, L. A., 1998, Present-Day Plate Boundary Digital Data Compilation: University of Texas Institute for Geophysics, 174.

Courtillot, V., Davaillie, A., Besse, J., and Stock, J., 2003, Three distinct types of hotspots in the Earth's mantle: Earth Planet. Sci. Lett., v. 205, no. 3-4, p. 295-308.

Della Mora, S., Boschi, L., Tackley, P. J., Nakagawa, T., and Giardini, D., 2011, Low seismic resolution cannot explain S/P decorrelation in the lower mantle: Geophys. Res. Lett., v. 38, no. L12303, p. 6.

Dingle, R. V., and Scrutton, R. A., 1974, Continental breakup and the development of Post-Paleozoic sedimantary basins around southern Africa: Geol. Soc. Am. Bull., v. 85, no. 9, p. 1467-1474.

Jaques, A. L., Lewis, J. D., and Smith, C. B., 1986, The kimberlites and lamproites of western Australia: Geological Survey of Western Australia, 132.

Jelsma, H., Barnett, W., Richards, S., and Lister, G., 2009, Tectonic setting of kimberlites: Lithos, v. 112S, p. 155-165.

Jelsma, H. A., de Wit, M., Theart, C., Dirks, P. H. G. M., Viola, G., Basson, I. J., and Anchar, E., 2004, Preferential distribution along transcontinental corridors of kimberlites and related rocks of southern Africa: South African Journal of Geology, v. 107, p. 301-324.

Kuo, B.-Y., Garnero, E. J., and Lay, T., 2000, Tomographic inversion of S-SKS times for shear velocity heterogeneity in D": Degree 12 and hybrid models: J. Geophys. Res., v. 105, no. B12, p. 28139-28157.

Kustowski, B., Ekström, G., and Dziewonski, A., 2008, Anisotropic shear-wave velocity structure of the Earth's mantle: A global model: J. Geophys. Res., v. 113, no. B6.

McMillan, I. K., 2003, Foraminiferally defined biostratigraphic episodes and sedimentary patterns of the Cretaceous drift succession (Early Barremian to Late Maastrichtian in seven basins of the South Aftican and sourhtern Namibian continental margin: South African Journal of Science, v. 99, p. 537-576.

Moore, A., Blenkinsop, T., and Cotterill, F. P. D., 2008, Controls on post-Gondwana alkaline volcanism in Southern Africa: Earth Planet. Sci. Lett., v. 268, p. 151-164.

-, 2009, Southern Africa topography and erosion history: plumes or plate tectonics: Terra Nova, v. 21, p. 310-315.

Morgan, W. J., and Morgan, J. P., 2007, Plate velocities in the hotspot reference frame, *in* Foulger, G. R., and Jurdy, D. M., eds., Plates, Plumes, and Planetary Processes, Volume Special Paper 430: Boulder, Clorado, Geological Society of America, p. 65-78.

Nürenberg, D., and Müller, R. D., 1991, The tectonic evolution of the south Atlantic from Late Jurassic to present: Tectonophys., v. 191, p. 27-53.

Reeves, C., and De Wit, M., 2000, Making ends meet in Gondwana: Retracing the transforms of the Indian Ocean and reconnecting continental shear zones: Terra Nova, v. 12, no. 6, p. 272-280.

Richards, M. A., Hager, B. H., and Sleep, N. H., 1988, Dynamically supported geoid highs over hotspots: Observations and theory: J. Geophys. Res., v. 93, no. B7, p. 7690-7708.

Ritsema, J., Heijst, H. J. v., Deuss, A., and Woodhouse, J. H., 2011, S40RTS: a degree-40 shear velocity model for the mantle from new Rayleigh wave dispersion, teleseismic traveltimes, and normal-mode splitting function measurements: Geophys. J. Int., v. 184, p. 1223–1236.

Romanowicz, B., and Gung, Y., 2002, Superplumes from the core-mantle boundary to the lithosphere; implications for heat flux: Science, v. 296, no. 5567, p. 513-516.

Savage, J. C., 1993, The Parkfield prediction fallacy: Bull. Seismol. Soc. Am., v. 83, p. 1-6.

Simmons, N. A., Forte, A., Boschi, L., and Grand, S., 2010, GyPSuM: A joint tomographic model of mantle density and seismic wave speeds: J. Geophys. Res., v. 115, no. B12310.

Sleep, N. H., 1990, Hotspots and mantle plumes: Some phenomenology: J. Geophys. Res., v. 95, p. 6715-6736.

Steinberger, B., 2000, Plumes in a convecting mantle: Models and observations for individual hotspots: Journal of Geophysical Research-Solid Earth, v. 105, no. B5, p. 11127-11152.

Steinberger, B., and O'Connell, R. J., 2000, Effects of mantle flow on hotspot motion, *in* Mark A. Richards, R. G. G., and Rob D. van der Hilst, ed., The History and Dynamics of Global Plate Motions: Washington, D. C., AGU, p. 377-398.

Thorne, M. S., Garnero, E. J., and Grand, S. P., 2004, Geographic correlation between hot spots and deep mantle lateral shear-wave velocity gradients: Phys. Earth Planet. Int., v. 146, p. 47-63.

Torsvik, T. H., Burke, K., Steinberger, B., Webb, S. J., and Ashwal, L. D., 2010, Diamonds sampled by plumes from the core-mantle boundary: Nature, v. 466, p. 352-355.

Torsvik, T. H., Smethurst, M. A., Burke, K., and Steinberger, B., 2006, Large igneous provinces generated from the margins of the large low-velocity provinces in the deep mantle: Geophys. J. Int., v. 167, p. 1447-1460.

Trampert, J., Deschamps, F., Resovsky, J., and Yuen, D., 2004, Probabilistic tomography maps chemical heterogeneities throughout the lower mantle: Science, v. 306, p. 853-856.