The Oil Drum
DISCUSSIONS ABOUT ENERGY AND OUR FUTURE

# Dispersion, Diversity, and Resilience

Posted by Gail the Actuary on February 27, 2010 - 10:28am
Topic: Environment/Sustainability
Tags: dispersion, diversity, george mobus, resilience, webhubbletelescope [list all tags]

*This is a guest post by WebHubbleTelescope.*

If we want to have any hope in controlling our destiny we have to understand our environment. In one sense, if we treat our environment as a control system, capable of responding to a stimulus, we need to understand not only its behavior, but how it will respond to the stimulus. One can ask: will it collapse in response to dwindling resources? Or will it rebound and stay resilient? For that we require a good model of the system. And of course, the simpler the model to describe, the better.

The system thinker and cyberneticist Ross Ashby summed it up with two seminal ideas. His simplicity criteria, the "Law of Requisite variety" states, "Variety absorbs variety, defines the minimum number of states necessary for a controller to control a system of a given number of states." Our capabilities thus become limited by the amount of information available to us. The second, the "Good Regulator theorem" goes "every good regulator of a system must be a model of that system". In other words, to regulate any causal system, we should require a model of how the system will behave normally and how it will react to a disturbance. The two relate directly to the classical view of control theory, that of controllability and observability. That may sound a tad idealistic, but that's how an engineer would respond to a problem statement.

The seeming diverse complexity of a system such as the Amazon rain forest, remains in many ways simple to describe. Interestingly, we can actually understand how that system evolves and adapts, given that we have a simple-enough model to work with. To compare and contrast, we can also consider a model of oil resources and how we can understand a seemingly random distribution of reserves.

Further, to have any chance of controlling the behavior we need good observability via good measurements. A human mobility metric is just one example of this, one simple to model, which gives us a good understanding, and one that we can monitor in the future.

For these three cases, (1) biodiversity, (2) oil abundance, and (3) human mobility, I will describe a few simple models based on entropy principles (maximum entropy dispersion, the "entroplet") and working with barest and most minimal information available to us. We will see how far that can take us. The discipline of complex and resilient systems remains wide-open for discussion.

---

George Mobus recently posted on Energy Flow, Emergent Complexity, and Collapse. As a response, I intentionally named this post Dispersion, Diversity, and Resilience. This doesn't mean that I disagree with his point-of-view, just that mine differs.

I have a long-running interest in the topic of disorder. When I first approach a problem, I try to characterize the observed behavior according to whether it follows a predictable, unpredictable, or random/noisy process. I rarely use the categories of complex or chaotic. In my mind, if you do that, you give up some hope in solving the problem.

So I tend to rebel against conventional notions of complexity. As defined by the scientific establishment, complexity seems to have taken on a discipline of its own. Once some problem gets characterized as complex or chaotic, the big thinkers emerge from the Nonlinear Institute of Profundity, leaving the rest of us behind. As a more pragmatic approach, I wouldn't mind capturing a complementary world-view to overtly sophisticated models of complexity. I notice that the expert-level mathematical explanations invariably become hairy, while the popular explanations lack common sense or excessively rely on heuristics (see oil depletion analysis as a primary example of the latter). However, the most elegant approaches tend to apply the simplest patterns to concisely describe the complexity. I elaborated on this topic in a previous post, so I won't try to explain the philosophy behind the "simplicity out of complexity" paradigm.

Instead, I want to practically demonstrate where and how some simple and rather parsimonious explanations can go a long way to explaining why disordered, yet seemingly complex, systems have very intuitive explanations. And in keeping with the intent of this post, I will try to show how these system might become resilient against collapse. I believe that we often mistake complexity for garden-variety disorder (i.e. entropy) and that the diversity that arises due to strictly entropic arguments may also demonstrate some robustness. In the end, if we can understand how the diversity arises, then we may figure out how to achieve or at least maintain a good outcome.

As a muddying factor, I sense that people's intuition fails when they try to deal with complexity or entropic disorder. Although I consider the supporting explanations and rationalization for disordered systems rather simple, they do build on some math. Ordinarily, this would get the math-phobes upset, and if this ever gets a wider audience, it may anger some of the experts on the subject as well. As you will see, the premise that I will outline essentially trivializes certain aspects of the "complex ecosystem". What we actually observe amounts to nothing more than our own filtered view on a disordered state. As the key point to retain, the observed disorder runs high enough so as to simplify the entire argument, both *mathematically* and *conceptually*.

In my previous post on crude complexity, I used the writings of Murray Gell-Mann to guide the narrative. He basically explained how seemingly complex systems often possess the simplest descriptions. To extend this context, I will use Edwin T. Jaynes research on probability and entropy to help justify the simplicity premise. According to Jaynes, we just scratch the surface of the practical applications of entropy if we consider it only as something that arises out of thermodynamics. Instead, if we treat entropy as a first-class measure of the disorder in a system, it can prove useful in many other scientific investigations.

Of course, we can make a connection to the oil depletion equation as well, as simplicity often proves contagious. The systems thinker John Gall (*Systemantics*) once said: *"A complex system that works is invariably found to have resulted from a simple system that worked"*.

I will go through a few cases of seeming complexity masquerading as random and disordered behavior. These share the approach of applying some basic ideas and common sense, and then working out the problems as you would in a chalk-talk. The shorter the better in that case, otherwise you risk having your audience doze off.

I tend to use the same math in all these analyses. So if you get stuck in understanding the principles in one of the explanations, something might spark your intuition in another. I apply the same math as a way to unify my understanding, as well as to substantiate the overall approach. As with most effective arguments, the more broadly we can apply the arguments, the more confidence we have in its generality and applicability. For many of the global problems we face, we don't have the benefit of a controlled experiment. The earth itself acts as both the test and the control. For that reason, if we can find unifying global behaviors, we gain confidence by the accumulation of these "proxy" explanations. As a side effect, you may end up finding quite a few interesting emergent results from the case studies.

## The first case:

**Relative Abundance Distribution**

In nature, the diversity of species gets reflected in the samples of various populations taken during scientific surveys. Scientists invariably find, and we get reminded quite often, that a few species predominate in their abundance while the majority of species have relatively sparse populations. Many species remain extremely rare or go undiscovered. This data typically gets plotted as a Relative Abundance Distribution (RAD) histogram.

As a result of these empirical observations -- usually taken in some very diverse populations of a certain category of wildlife or plant -- you will actually see relative counts of the most common species outnumber the rarest species by orders of magnitude. In sampling experiments, the rarest species may actually have counts of only 1.

Most recent research on this topic has concentrated on understanding the relative species abundance (RSA) of somewhat isolated ecosystems. Understandably, scientists approach it this way so that they can limit or control the set of measurable parameters and therefore understand the phenomena on a larger and more heterogeneous scale. This 2007 article in **Nature**, Patterns of relative species abundance in rainforests and coral reefs, suggests that *interactions* among species don't have as large an effect as imagined.

> **Abstract**: A formidable many-body problem in ecology is to understand the complex of factors controlling patterns of relative species abundance (RSA) in communities of interacting species. Unlike many problems in physics, the nature of the interactions in ecological communities is not completely known. Although most contemporary theories in ecology start with the basic premise that species interact, here we show that a theory in which all interspecific interactions are turned off leads to analytical results that are in agreement with RSA data from tropical forests and coral reefs. The assumption of non-interacting species leads to a sampling theory for the RSA that yields a simple approximation at large scales to the exact theory. Our results show that one can make significant theoretical progress in ecology by assuming that the effective interactions among species are weak in the stationary states in species-rich communities such as tropical forests and coral reefs.

I have no problems with their assertions, only that the math that they invoke goes a bit overboard, and provides very little insight. Practically speaking, with just a few nods to maximum entropy, we can show agreement to the results in a few lines of derivation. With that simplicity, we get the benefits of a significant amount of extra insight. Also note that the authors state that they don't know the interactions; this uncertainty suggest that entropy arguments may work out well. In other words, maximum entropy provides an avenue for reasoning about an uncertain world.

To derive a universal RAD, we start out with a few assumptions.

We first assume that different species evolve as random processes that essentially fill up space. I propose a quantity that, for the lack of a better term, I call the adaptation level, *A*. This can have the units of, for example, **#organisms/acre** so it maps to a parameter proportional to sampling some species.

We next use the Maximum Entropy Principle (MaxEnt) to describe the uncertainty in the time it takes for a species to reach an adaptation level.

> $p(t) = (1/t0)\, \exp(-t/t0)$
> $P(A,r \mid t0)$ = integral of $p(t)$ for all $t$ such that $t$ is greater than $A/r$

These are both probability distributions, the first a density function and the second a conditional cumulative probability.

The first expression essentially states that we don't know how long it takes to reach a given adaptation level, only that it has a mean time, *t0*. This mean time could reach millions of years, but for now the specific value doesn't matter. We just assume that the likelihood of times around that mean has a maximum entropy described by the exponential probability function, *p(t)*. We then assume a rate, *r*, that relates a time to reach a given adaptation level, $A = r*t$. Every species that reaches the adaptation level has to evolve for a time $t=A/r$, so the conditional cumulative probability is described as $P(A,r \mid t0)$ and derives to the value below.

> $P(A,r \mid t0) = \exp(-A/(r*t0))$

We also have uncertainty in the adaptation level, *A*, assuming only that it also has some mean value *A0* with the same MaxEnt probability density function.

> $P(r \mid t0,A0)$ = integral of $P(A,r \mid t0)$ over all $A$

This results in the cumulative adaptation rate function as follows:

> $P(r \mid t0,A0) = 1/(1+A0/(r*t0)) = 1/(1+r0/r)$

If we replace *A0/t0* with *r0*, we see that the above relation describes a set of species that evolve with a huge dispersion -- high enough for it to describe a fat-tail distribution of rates. In other words, the large disorder in both the time scale and adaptation level generate a rate function that generates an even larger entropy in the evolution of various species. The two degrees of freedom in uncertainty gives it *double* the entropy of a single MaxEnt exponential probability density function. This uncertainty results in a very disordered system. I make the claim that diversification and growth of speciation possesses maximum entropy. However consumed energy aids in driving adaptation, it doesn't affect the probabilities, as entropy plays the lead role in generating the dispersion.

The relative abundance comes about when you consider that low values of *r* will lead to smaller relative population levels than higher values of *r*. Since probabilities get invoked, we also see the effects of abundance as a combination of population size and rarity. In other words, abundance essentially relates the value of *r* (proportional to the size of the species population when evaluated over a period of time) to rarity, which states how often that size of population occurs. (To foreshadow a bit, the same argument holds in sizing oil reservoirs.)

At this point, I don't necessarily care how the authors' of the **Nature** article derived their own model, just that it has greater informational complexity than my derivation. Since the equation for $P(r \mid r0)$ contains only a *single adjustable parameter*, *r0*, it meets the Gell-Mann acid test for simplicity. When we transform *r0* to an abundance, we use a proportionality constant, *k*, and call the result an abundance dispersion factor, $D=k*r0*time$.

> **CDF** : $P(X) = 1/(1+D/X)$
> **PDF** : $p(X) = D/(D+X)^2$

The basic normalized shape looks like the following histogram plotted on a double logarithmic scale (the log-normal heuristic is shown for comparison):

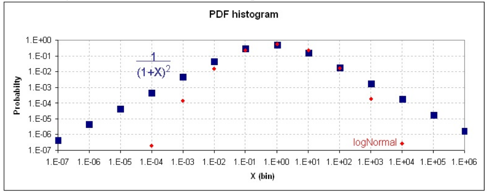


**Figure 0**: Probability Density Function (PDF) histogram of the entropic dispersion function, normalized to 1. Plotted per decade on a logarithmic scale, the function appears symmetric. The function is so ubiquitous and has such nice aggregating properties, I will call it an *entroplet*.

On this scale $D$=1, and the term $X$=abundance corresponds to the relative size of the population, while *p(X)* provides the abundance of that population in terms of a probability. Therefore, you can read it as either very small or very large *X* populations occur infrequently, with the peak frequency lying in between the extremes. However, since large populations consist of large numbers of organisms, they do show up more often in statistical samples. The rarest species may never show up in samples (the so-called "Preston's veil") both because rare species have a small population and because that size population rarely occurs over the epochal time integration considered. Interestingly, no mean value exists for this PDF, which often happens for fat-tail distributions. In practice, this has little impact in a finite world, as we consider both finite time and space to generate limiting constraints.

That essentially describes the extent of the theory. Simple enough, so let's see how effectively a single parameter fit works with the observed data.

**Results.** Data of relative abundance usually gets plotted as a cumulative rank histogram (also known as a Whittaker plot) or as a Preston plot (which essentially describes the probability density function (PDF) as a logarithmically binned histogram).

I took data from the **Nature** article and applied the simple theory on Whittaker histograms first. The following diagrams contain a single-parameter fit to the data, shown as the RED lines. Each diagram corresponds to a different isolated tropical forest region and the RAD for sampled tree species within those regions. The BCI region has a dispersion factor of $D$=23 while the Pasoh region has a factor of $D$=14. Since this function has scale-free properties, the dispersion factor really only shifts the location of the knee in the curve along the abundance axis. The BCI has a higher mode for the relative abundance than the Pasoh region, which could imply that the most common species adapted faster for the BCI region or that the BCI region evolved/adapted over a longer period of time. Importantly, we can't tell the difference because we have derived *A0/t0* only as a ratio; in other words, we have lost the ability to separate the two effects.

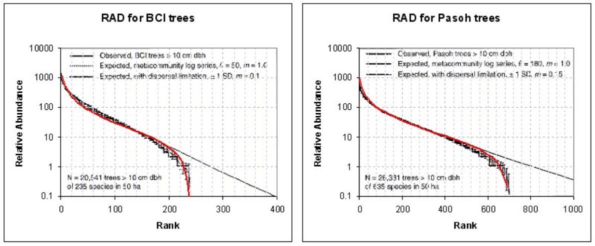


**Figure 1**: RAD histograms for isolated adaptation regions. The red lines shows the dispersive model.

The movement of the mode becomes more apparent if graphed as a Preston plot. This bins the data on a more granular level, yet gives a view that more intuitively shows the most frequent population as a peak value. The same data shows up as BLUE dots below.

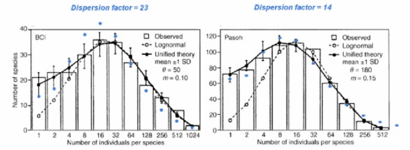


**Figure 2**: RAD histograms as Preston plots show the mode position.

The general trend shows some universality. Plotted below in RED dots is the fit to a sampling of moth populations.

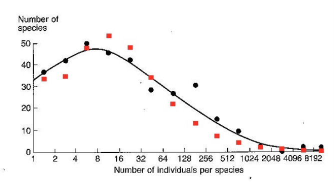
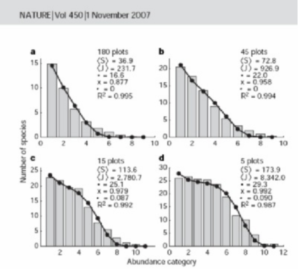


**Figure 3**: RAD histogram for moth population. Red dots show the dispersive model.

Remember, the factor *D* sets the peak position and provides the only adjustable parameter in the fit. Maximum entropy considerations alone set the width of the curve. For the moth RAD histogram shown above, the black line gives the log-normal fit. This may appear better than the dispersive fit, but the log-normal has **three** adjustable parameters available for tuning, and has little additional intuitive significance.



Not all of the RAD plots show a mode peak away from low abundance. For example, the adjacent figure reproduced from the **Nature** article generates data for some localized coral communities. From what I understand, these have gotten somewhat isolated from the larger metacommunity which understandably would show a larger diversity in the coral population. The isolation gives many low abundance species, binned as base 2 abundance categories between 0, 1 and 2 ($2^0=1$, $2^1=2$, $2^2=4$).

Qualitatively this also agrees with what the dispersive formulation says if we apply low values of the dispersion factor, *D*. Again, these populations may not have evolved/adapted as fast or they may not have evolved/adapted over a long enough period. Without additional information, this is all that the maximum entropy principle can tell us.

If we plot a range of dispersion factors on the same binned max rank of samples, the results appear as below. Apart from low values of *D*, varying the value of *D* simply shifts the distributions away from the origin. Each shape defines an *entroplet*, which lacks any parameter besides its position, becoming essentially scale-free. On the horizontal logarithmic scale, the width looks like it remains constant, but it actually spreads out to accommodate a larger range of individuals per species.

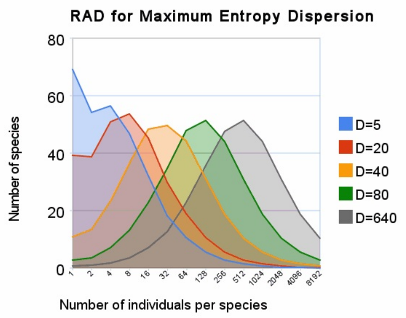


**Figure 4**: Dispersive model with various values of D

This article further describes the extent of the metacommunity diversity. The histogram on the left below shows the dispersion of coral species on a more local scale while the plot on the right duplicates that result while also displaying the large shift (256 much larger than 66) on a metacommunity scale. So even though the dispersion factor changes by nearly a factor of four, the range in number of individuals per species agrees perfectly with the maximum entropy formulation.

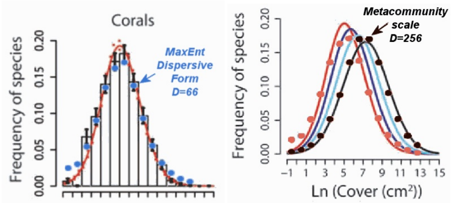


**Figure 5**: RAD histograms on coral communities. MaxEnt dispersion model is a single parameter

fit.

The dispersive formulation will likely work on any general population simply because the relative abundance of species results from a principle of maximum disorder in the adaptation rates. With that uncertainty in place, the Maximum Entropy Principle guides us to the correct distribution.

This brings up an interesting situation. What happens when we try to apply this construct on a massively larger metacommunity scale? The tree diversity of the entire Amazon basin provides a situation that we can analyze in context. The distinction here is that the heterogeneous nature of the geography and geological events places the dispersion on a smeared time-scale. As you can see below, the histograms show a much broader shoulder than the isolated adaptation results.

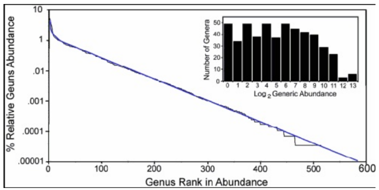


**Figure 6**: RAD histograms of trees in the Amazon basin. (Ignore the jump between 1 and .01 on the y-scale, as this looks like a typo)

To model the effect on a larger scale using MaxEnt dispersion, we have to consider a spread in time ranges. Unlike isolated regions, such as might happen on an island (e.g. BCI = Barro Colorado Island), adaptation did not start at one specific time in prior ecological history. Instead, due to a variety of factors, which can include mass extinctions, introductions, and climate change, the effective start time for adapation ranges over a scale aligned with historical events.

The time integration runs from the first significant event at a time $t+T$ ago to the last significant event $t$ ago (also the maximum entropy result for a range). The value of $T$ denotes a very large ecological time scale in comparison to $t$.

P(X) = Time integral from $t$ to $t+T$ of *k*r0*time/(D+k*r0*time)^2*
P(X) = C*ln( (1+X/D1)/(1+X/D2) )

This generates a logarithmic-shaped function that has a much more diverse spread in abundance levels compared to the isolated adaptation result. The values of $D1$ and $D2$ correspond to an epochal time range, and $C$ is a constant that normalizes the result to 1 as X goes to infinity. Compare the red line below to **Figure 6**.

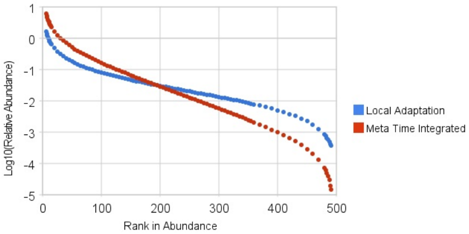


**Figure 7**: RAD histograms of dispersive models with specific Local start time and smeared Meta start time. For Meta, $D$ ranges from D1=1.7 to D2=3300. For Local, $D$ is fixed.

If you look at **Figure 4**, you can almost intuit what happened during the time integration. Each of the peak positions for various values of $D$ gets spread out over time so that the linear superposition of the individual curves creates a broad almost flat-topped peak on the binned histogram as shown below. The interpretation for this uniformity makes intuitive sense: if you reach far back enough, every effective adaptation level would be achieved, split equally between each doubling octave. The effects of dispersion spreads out the impact over time so that each generation contributes evenly. In my book, that defines an existential entropy, conceptually similar to the flat white noise spectrum that forms the background radiation in the universe -- in other words, noise as a form of ultimate disorder.

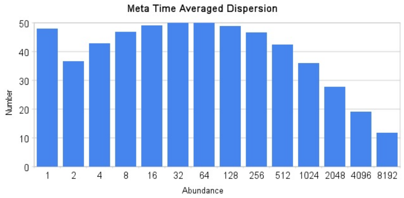


**Figure 8**: RAD histogram of time integrated metacommunity. Compare with the inset of Figure 6.

That basically explains the diversity of species according to entropic dispersion. Resilience occurs as a result of the shear diversity in the species. If one species becomes extinct, another one will likely take its place. It make take a long time, but it certainly will happen.

If you want to get really overloaded in math, go to the papers written by the authors of the **Nature** article. They don't seem to mention entropy at all, preferring to use conventional combinatorial statistical arguments (exactly how you can derive entropy, by the way). So you can take the long way around the lake (their approach) or take the scientific shortcut across the water (my approach). I essentially ignored their overly sophisticated derivation because it lacked the simplicity that these problems *should* demonstrate. Go with the math that I have described and you can actually try out the models yourself and get the insight that I have reached. Recall that this model contains only a single parameter and, quite fundamentally, we have nothing to fear as the results derive from garden-variety disorder and the complexity essentially wilts away.

I especially harp on the math model because the math alone shows significant similarity to what happens when we consider the sizing and abundance of oil reservoirs. That serves as our second case.

## The second case:

**Oil Reservoir Size Distributions**

I use essentially the same entropic dispersion formulation to describe the variation of reservoir sizes in the context of oil exploration. Instead of searching for living organisms and ranking the relative abundance, we sample geological formations and rank order the sizes of reservoirs we find. I go into the details of this approach in previous posts so won't repeat the details here. The fact that a dispersion form works just as well for oil as it does for species has to do with the disordered range in rates that go into reservoir formation. The figure below shows the agreement for entropic dispersion for North Sea reservoir data.

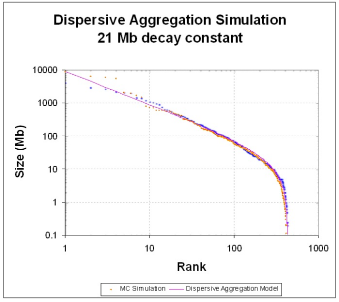


**Figure 9**: Data of reservoir sizes for North Sea (blue dots) plotted alongside dispersive model.

I find it intriguing the similarity between the the large population of a few species of living organisms and the large size of just a few oil reservoirs. The same MaxEnt math generates exactly the same fat-tail distributions.

Another interesting analogy in reservoir size distribution relates to how the dispersion factor varies significantly in different regions of the world. In particular, the USA has a dispersion factor that appears lower than elsewhere. Fitting the values for the USA, we see a value of around 1, whereas the North Sea has a value of 21 (see Figure 8).

**Figure 10**: Dispersion of reservoir sizes for USA. Time aggregation of entropic dispersion does not affect the tails of the distribution, only the shape of the knee, making it less sharp.

I use the same interpretation here as I use in species adaptation. If I assume that the isolated North Sea region "evolved" from a point in time long ago, then the single entroplet fits the data well. However, the entire USA shows a much more heterogeneous nature, partly due to its geographic area, and we can use a maximum entropy estimator to uniformly spread the entroplet functions over a range in start times (i.e. the MEP estimator for a fixed range is a uniform density). So if we apply this to the USA, the fit becomes better and the dispersion factor increases to better match that observed in other parts of the world (usually between 10 and 30). In other words, the North Sea acts as a localized community and the USA provides a metacommunity in the analogous ecosystem sense. The solid black line overlays the data points; we could interpret the range as occurring between 500 million years ago and 0.5 million years ago, with anything shorter than this time spilling beyond the MaxRank of 14,000 fields (field data from 1986).

The flattening of the PDF that occurs in the Amazonian case for tree diversity also occurs in the USA for reservoir size diversity. The uniform time-shifted aggregation of the entroplets does not affect the tails but makes the mid-range of sizes more equally abundant. As a bit of insight, this likely accounts for the greater percentage of low volume stripper wells in the USA.

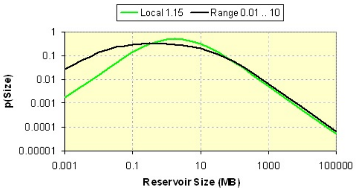


**Figure 11**: PDF of an reservoir sizing entroplet, and the time-shifted aggregation of a range. The knee in Figure 10 becomes a flatter top, but the fat-tails remain unaffected.

I suppose that the the USA could have expended more effort in finding small reservoirs than other oil-producing countries, which may contribute to a lower range of values for *D*. The geological rate of reservoir formation also could have progressed on average much more slowly here than elsewhere. Or, more likely according to this model, it has started from a more recent geological epoch, so it hasn't matured as much as the rest of the world. Whether this has consequences, or remains an inconsequential data point will have to wait until we get better data.

In a sense, the diversity of reservoir sizes around the world has effectively reached that of a set of biological species; the origin of oil formation in geological terms occurred long ago, but geological activity has likely allowed the movement of oil to effectively "restart" many times over epochs. Since I showed that the greater diversity of USA oil follows the pattern of Amazonian tree species, we should also find a more uniform density of oil reservoir sizes around the knee of the curve across the global spectrum.

However, unlike the resilience of a diverse biological population, no rebound effect occurs for oil depletion. Once we deplete oil, it doesn't come back. The size of the reservoir doesn't matter. Other smaller size reservoirs can take up the slack, but unlike the diversity of living organisms, dead organisms do not recover.

Many other analogs exist between species search and reservoir discovery. For example, shown below is the species equivalent of a creaming curve (from Species Abundance Patterns). The rate at which we find oil reservoirs has a close analogy to the rate at which we find species. In terms of the reserve gowth issue, this has the same uncertainty in knowing when you have reached an asymptote..

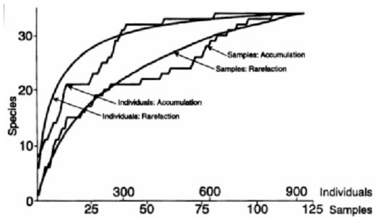


**Figure 12**: Species equivalent of a creaming curve.

## The third case:

**Distribution of Human Travel**

A final analogy has a purely human element. Human travel and mobility patterns in the USA show the same dispersion formulation with excellent agreement to recent data. We might imagine that human travel patterns would follow some complex behavior, yet if we simply assume that the *delta X* position changes and the *delta T* time changes each follow maximum entropy probability density functions, then the travel patterns reduce to a simple dispersive result (shown as the blue dots below).

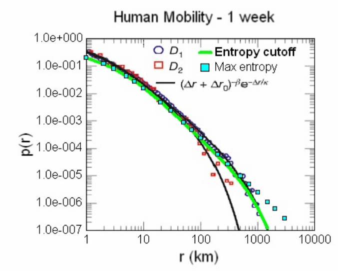


**Figure 13**: Human mobility in terms of dispersive model.

The green curve generates a single parameter fit to measured cell-phone usage data. The expression shown in the inset is a heuristic developed by the authors of the original Nature article in which they describe a "magic" exponent, beta. That fits equally well but has no basis in fundamental understanding. The beta term shown happens to come close (1.75+/-0.15) with the entropic dispersion exponent of 2. Unfortunately, the academic discussions surrounding human mobility appear even more sophisticated than that for eco-system diversity and the simple understanding gets lost.

As an interesting experiment, the dispersive function can be generated via a Monte Carlo (MC) simulation drawing from two MaxEnt variants, delta A and delta T, and then dividing the two, resulting in a set of sampled rates. One such MC run looks like the following, with the analytical result overlaid. In the real world, the effect of reduced sample space generates the noise observed.

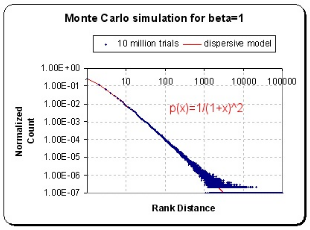


**Figure 14**: Generating dispersion rates by deltaX/deltaT MaxEnt variates.

The resilience to human travel patterns in the face of dwindling oil supplies will become important in the future. In terms of the dispersive model, the value of D will likely shift to smaller values without changing the nature of the curve. We will still live in an entropic world, but the energy that allows us to move around easily will inhibit our resiliency.

## Discussion

The simple theory defined here certainly does not qualify as blind curve fitting. Although abstractly defined, the assumptions follow from scientifically valid premises. As Jaynes suggests, you use maximum entropy arguments when you face any degree of uncertainty in your numbers. If you have better numbers, you can use those. The Maximum Entropy Principle has a close relative in Bayes Rule in this respect. As it stands, like Bayes, the model helps to manage our uncertainty by including valid prior information.

**Dispersion** : The effects of disorder result in variation of adaptation rates as a form of dispersion. This has more to do with entropy than energy flow.
**Diversity** : The huge changes in relative abundance comes about from the dispersion. I do not consider this emergent complexity, instead it describes diversity predicated on disorder. One man's perceived complexity is another's effective simplicity.
**Resilience** : The significant diversity derived simply from disorder considerations leads to the possibility of resilience against potential collapse. Since species may not have as much interdependence as assumed, it seems intuitive that the diversity can act as a buffer against extinctions. If one species becomes extinct, another more slowly evolving species may take over.

Using the same arguments as for species adaptation, we can arrive at similar results for reservoir sizing. The rather simple logical arguments should prove useful in any analytical context that proceeds under disordered, entropic conditions (see **case 3** for a largely sociological example). In this regime you can't use deterministic models such as Lotka-Volterra, and you need instead to consider probabilities for all your measures. Jaynes had it right when he titled his final book as "**Probability Theory: The Logic of Science**".

Besides Jaynes, the pioneers of fat-tail and fractal statistics have contributed some interesting insight, especially in pointing out that fat-tail statistics occur more often than common wisom dictates. Taleb's admonition to not overuse Gaussian/normal statistics becomes very important when working under maximum uncertainty. As an example, the fact that we know the variance of a process, would suggest that we use a normal distribution, yet we have no knowledge of the standard deviation of any of our data sets. We barely have knowledge of the mean as it stands.

A bit puzzling in that I can imagine that Mandelbrot could likely have derived the dispersive model as it has a remarkable similarity to the discrete power-law Zipf-Mandelbrot heuristic. Mouillot (reference below) mentions Zipf-Mandelbrot in an abstract with reference to RAD but I do not have access to the full document. Mandelbrot, and Taleb for that matter, appear to show some disdain for actually deriving any of these distributions, preferring instead to describe fractal or fat-tail behavior as heuristic functions. Although they have very good insight regarding uncertainty and randomness, they seem to prefer the world of descriptive rather than proscriptive statistics. Many of the diversity research efforts appear to think that modelling only involves heuristic fits of previously categorized statistical distributions to the data (reference here) . Only the authors of the Nature article have a model (the Neutral theory) to base predictions on, yet that model will likely prove too intricate to get burned into our consciousness.

Why no one else has previously formulated such a simple model as maximum entropy dispersion would require a separate discussion. It may exist somewhere in the literature but I have yet to find it. I realize that my arguments have some abstract concepts, such as equating uncertainty with entropy, that may take some getting use to. Suffice to say, Taleb may have captured the :

> *"We scorn the abstract, we scorn it with a passion."* from **The Black Swan**

As any software developer understands, the right abstraction often helps to clean up complexity. It just takes the right insight to get to that abstraction.

---

**References**

1. Igor Volkov, Jayanth R. Banavar, Stephen P. Hubbell and Amos Maritan, *Neutral theory and relative species abundance in ecology*, **Nature 424**, 1035-1037 (28 August 2003) http://arxiv.org/PS_cache/q-bio/pdf/0504/0504018v1.pdf
2. *Differential Entropy and Probability Distributions* http://www.mtm.ufsc.br/~taneja/book/node14.html
3. *Towards a general theory of biodiversity* http://www.nature.com/nature/journal/v410/n6831/full/410923a0.html

4. *A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs*
   http://sciwebserver.science.mcmaster.ca/biology/faculty/kolasa/Hubbell_1997.pdf
5. Sean R. Connolly, Terry P. Hughes, David R. Bellwood, Ronald H. Karlson, *Community Structure of Corals and Reef Fishes at Multiple Scales*, **SCIENCE** VOL 309 26 AUGUST 2005
   http://www.reefresilience.org/pdf/1363-Community_Structure_of_Corals_and_Reef_Fishes_at_Multiple_Scales.pdf
6. *Species Abundance Patterns*
   http://www.life.illinois.edu/ib/453/453lec16spprichness.pdf
7. *The Theory of Island Biogeography*
   http://web2.uwindsor.ca/courses/biology/macisaac/55-437/lecture9.htm
8. "*Extinction most likely for rare trees in the Amazon rainforest*"
   http://www.eurekalert.org/pub_releases/2008-08/stri-eml081308.php
9. MOUILLOT D., LEPRETRE A., *Introduction of relative abundance distribution (RAD) indices, estimated from the rank-frequency diagrams (RFD), to assess changes in community diversity*, **Environmental monitoring and assessment**, v.63, 2000.
10. Stephen P. Hubbell, Fangliang He, Richard Condit, Luis Borda-de-Agua, James Kellner, and Hans ter Steege, "*How many tree species are there in the Amazon and how many of them will go extinct?*", **PNAS** August 12, 2008 vol. 105 suppl. 1
    http://www3.bio.uu.nl/peb/staff/personal/htersteege/pdfs/Hubbell%20et%20al%202008%20How%20many%20tree%20species%20are%20there%20in%20the%20Amazon%20and%20how%20many%20of%20them%20will%20go%20extinct%20PNAS%20105.pdf
11. http://en.wikipedia.org/wiki/Niche_apportionment_models
12. *Simulated geographic variations of plant species richness, evenness and abundance using climatic constraints on plant functional diversity exponential curve describing rates.*
    http://www.bgc-jena.mpg.de/bgc-theory/index.php/Pubs/2009-ERL-AK-JA-RP-BR
13. *The model also has similarities to the Principle of Least Effect, which has both physics and information science origins.*