



# The Thermodynamic Theory of Ecology

Insights from information theory illuminate nature's large-scale patterns.

By Veronique Greenwood



The Western Ghats in India rise like a wall between the Arabian Sea and the heart of the

subcontinent to the east. The 1,000-mile-long chain of coastal mountains is dense with lush rainforest and grasslands, and each year, clouds bearing monsoon rains blow in from the southwest and break against the mountains' flanks, unloading water that helps make them hospitable to numerous spectacular and endangered species. The Western Ghats are one of the most biodiverse places on the planet. They were also the first testing ground of an unusual new theory in ecology that applies insights from physics to the study of the environment.

[John Harte](#), a professor of ecology at the University of California, Berkeley, has a wry, wizened face and green eyes that light up when he describes his latest work. He has developed what he calls the maximum entropy (MaxEnt) theory of ecology, which may offer a solution to a long-standing problem in ecology: how to calculate the total number of species in an ecosystem, as well as other important numbers, based on extremely limited information — which is all that ecologists, no matter how many years they spend in the field, ever have. Five years ago, the Ghats convinced him that what he thought was possible from back-of-the-envelope calculations could work in the real world. He and his colleagues will soon publish the results of a study that estimates the number of insect and tree species living in a tropical forest in Panama. The paper will also suggest how MaxEnt could give species estimates in the Amazon, a swath of more than 2 million square miles of land that is notoriously difficult to survey.



John

Harte thinks it is possible to predict the behavior of ecosystems using just a few key attributes. His method ignores nature's small-grained complexities, which makes many ecologists skeptical of the project.

If the MaxEnt theory of ecology can give good estimates in a wide variety of scenarios, it could help answer the many questions that revolve around how species are spread across the landscape, such as how many would be lost if a forest were cleared, how to design wildlife preserves that keep species intact, or how many rarely seen species might be hiding in a given area. Perhaps more importantly, the theory hints at a unified way of thinking about ecology — as a system that can be described with just a few variables, with all the complexity of life built on top.

## The Big Picture

Harte has an impressive track record as an ecologist. But before he entered the field, he was trained as a theoretical physicist. In his first faculty job 46 years ago, he taught thermodynamics at Yale University. “That’s when I first really became enamored of the foundations of thermodynamics and statistical physics, when I realized the power of the ideas that those theories are based on,” he said. In particular, he was fascinated by the idea that you could look at, say, a container of hydrogen and infer micro values, like the velocities of the molecules, from macro values like temperature and volume.

But he soon left to become an ecologist, studying the effect of acid rain on salamanders. Twenty-five years ago he began a landmark experiment at the Rocky Mountain Biological Laboratory, gradually heating up a subalpine meadow using electric heat lamps to simulate the climate of 2050 in order to discover what it will do to the soil and organisms found there. Thirty papers and nine doctoral theses have come out of the experiment, which is still running. “It’s been a major preoccupation of mine for a quarter of a century,” he said. About 15 years ago, however, he grew interested in macroecology, which deals with the search for large-scale patterns in ecosystems.

Ecologists study the connections between species and their environment, traditionally through detailed observations of the natural world. They might penetrate far into a rainforest, learning the calls of birds one by one until they identify one they’ve never heard before. They might, as Harte does, monitor a single meadow for decades, becoming deeply versed in the details of each creature’s existence. Many are also interested in high-level, abstract questions, such as how birds first began to flock. But the field is rooted in a kind of natural history.



Ecologists often come to understand an ecosystem by conducting arduous surveys, in which they record every plant or animal in a set of plots.

Macroecology deals with patterns that might be universal across ecosystems. When the field arose in the 1970s, ecologists tried to model the environment as a well-oiled machine that, given enough time, would settle into certain patterns. Yet when it became clear how much messier the real world is than those models, the field went quiet. “We were trying to answer bigger questions than our data could support,” said [William Kunin](#), a professor of ecology at University of Leeds in the U.K. who

watched the field evolve as an undergraduate in the 1970s.

In the late 1990s and early 2000s, macroecology rose again, driven by the need to understand the effects of mass deforestation, climate change and other large-scale changes in the environment.

"We're in a situation where there are big global-scale trends in species distributions, in climates, in fertilization of the planet. We're doing big things to the world," said Kunin, who now does

macroecology work. "And policymakers want from us answers of what that will do to biodiversity."

[Vanessa Weinberger](#), a doctoral student at the Pontifical Catholic University of Chile who has interned with Harte, adds: "What these people started to do was to try to come up with laws of ecology."

In Harte's first foray into this area, working from [a paper by Kunin](#), he asked whether species abundance could be structured like a fractal — whether the abundance would remain the same no matter what scale was used. He published a number of predictions that he tested with data from the real world. The results were unequivocal: "Nature is not fractal. It was wrong," he said.

Harte went back to thinking about thermodynamics, until he learned about a procedure called maximizing information entropy that was developed in the mid-20th century by information theorists who had been inspired by thermodynamics. Using this tool, he could start with a macro quality, the number of species counted by field ecologists in a 50-hectare plot of forest, and predict relatively micro qualities, such as how species would be distributed across much smaller subplots. After completing [that work](#), he had a pivotal conversation with his brother. "He said, 'You're scaling down. Can you run it backwards? I bet you'll run into trouble.' It took me an afternoon," Harte recalls, "and I figured out how to run it backwards." Harte had found a way to estimate species richness at much larger scales than he could measure.

"Once I realized you could do it, I realized there's no reason to stop at 50 hectares. You can go up, up, up," he said. And in the same moment, he thought of the Western Ghats. He had played with data on tree species in those mountains for the fractal theory. The data was unique because ecologists already knew approximately how many tree species there are — around 1,000 — but all previous attempts to scale up to the total number from plot data had underestimated the true count by 400 to 500. With MaxEnt, Harte got a result of 1,070. In 2009 he published [a paper reporting those results](#).

## Best Guesses

Places like the Ghats hammer home how complex ecosystems can be. They host a dizzying array of organisms — dragonflies, tigers, beetles, chameleons, 15-foot-long cobras, rare lion-tailed macaques, and the largest population of Asian elephants in the world. Current methods of quantifying such biodiversity are hit-or-miss, however, because no one yet understands how much overlap there is between what can be measured in a small survey plot and what else might be out there. Most estimates are lower than reality, researchers think.



When researchers find a particular species in a survey plot, they must then estimate how common that species is across a larger swatch of land. MaxEnt gives them a way to do that.

The power of the MaxEnt approach to ecology is that it doesn't try to deal with the details of these species — it doesn't take into account, for instance, how fast an elephant moves, or whether macaques are territorial. Instead it uses ideas from information theory to calculate the likeliest scenarios that could have given rise to fragmented data.

For example, imagine that you're shown a bag and told that it contains both red and blue marbles. Your task is to draw a few marbles out of the bag and, based on this limited information, estimate what fraction of marbles in the bag are red. (Researchers call this the probability distribution of marbles.)

Clearly, the more marbles you draw, the better your estimate will be. But in the context of ecology — where your bag might be the Amazon and the marbles insects — you'll only ever be able to pick out a tiny fraction of the total. MaxEnt is a statistical procedure that lets researchers estimate the most likely probability distribution of marbles in an ecosystem-size bag.

[Steve Pressé](#), a biophysicist at Indiana University-Purdue University Indianapolis who wrote a recent review that describes the technique's history, explains that MaxEnt is a way to make sure that conclusions drawn from small amounts of information are logically consistent. "The point is that if I have limited data, then in principle there's a whole bunch of probability distributions that could be consistent with that data," Pressé said. "How do I select the best one, the optimal probability distribution, given my problem?" MaxEnt is a way to do that.

Harte says he usually gets two responses: "You've opened up a whole new theory, and you're an idiot."

## Simple or Simplistic?

MaxEnt is based on principles of simplicity and consistency, but it has additional assumptions baked into it, starting with the fact that researchers must choose just a few variables to feed into the

procedure. In 2008, when Harte first considered the idea, he decided to try it out using the size of an area, the number of species there, the number of individuals, and the total metabolic rate of all those organisms. He didn't pick these characteristics at random; he had an inkling, from reading work on metabolic theory, that these had promise for describing biological systems. In some cases, they do very well.

The simplification of a complex ecosystem into just a handful of variables has fueled criticisms of MaxEnt, because it assumes that those numbers and whatever processes generate them are the only things shaping the environment. In essence, it generates predictions of biodiversity without taking into account how that diversity arises. It implies that the details many ecologists focus on might not matter if you want to understand the larger patterns of an ecosystem. Harte said he usually gets two responses: "You've opened up a whole new theory, and you're an idiot, because we all know that mechanism matters in ecology."

It's a difficulty many people have with models coming out of macroecology, including one of the first and best known, proposed by [Stephen Hubbell](#), an ecologist at the University of California, Los Angeles. Hubbell's neutral theory shows individuals living, dying and replacing each other, eventually generating an outcome that, as with MaxEnt, can look a lot like reality. Even though Harte, Hubbell and others [present their work as null models](#), which show what would happen in the absence of any other important processes, their success makes some ecologists uncomfortable.

"MaxEnt breaks everything that we ever thought about communities and species and ecology," Weinberger said. "If you go to a community, you're going to find a billion parameters. Ecologists are going to say, 'You have to take care of the wind, you have to take care of the water, and what if the lion has a headache?' They are going to try to measure everything. And that's the cool thing about John Harte: He is saying, let's keep it simple ... and just with four parameters, at a mini or a macro scale, you can figure out how to describe these patterns."

## Where It Breaks

With the Ghats, Harte was able to scale up from survey plots of a quarter of a hectare to the range's full area of 60,000 square kilometers. In his latest, still unpublished work, he and colleagues use survey data from a dozen 0.04-hectare plots in the San Lorenzo Protected Area forest in Panama, [published in Science in 2012](#), to generate an estimate of arthropod species on the level of the whole forest, about 6,000 hectares. Harte's new work also uses limited survey data to calculate the total number of arthropod species in the Amazon.

The authors of the Science paper used several other extrapolation methods to generate their own numbers in Panama, which came out significantly lower than Harte's. That discrepancy is interesting, said [Vojtech Novotny](#), an entomologist at the Biology Centre of the Academy of Sciences of the Czech Republic and one of the authors of the Science paper. "We don't know whether [MaxEnt] is an improvement or not," he said. The only way to know for sure is to gather more data. "[But] I see it as a very useful addition to the general exploration of extrapolation methods."

Because of the assumptions that restrict it, MaxEnt doesn't always work. For example, it breaks down in ecosystems that are undergoing rapid change. Kunin suggests that this failure exposes one of the theory's weaknesses as a predictive tool — the whole world is changing, and places that are stable aren't so easy to find anymore.

To understand why this failure happens, and to come up with a fix, Harte and his collaborators have taken censuses of insect species across the Hawaiian archipelago, from the freshest islands, just recently risen out of the ocean and still being colonized by life for the first time, to the oldest. He has

---

also obtained data from ecologists surveying the area around Cape Town, South Africa, a location that includes a mixture of disturbed and undisturbed plots.

With this information, Harte hopes to take the next step toward understanding when MaxEnt works and when it doesn't. The theory's breakdown may even be a useful marker of whether an ecosystem is disturbed. "Its value," Harte wrote in 2008, "derives in part from the nature of its failures."

*This article was reprinted on [Wired.com](#).*