



The real butterfly effect and maggoty apples

Tim Palmer

EVEN THOUGH THE NAVIER–STOKES
EQUATIONS ARE DETERMINISTIC, IT SEEMS
THAT YOU CANNOT MAKE PREDICTIONS
BEYOND A FIXED TIME HORIZON, NO MATTER
HOW SMALL THE INITIAL UNCERTAINTY.



Tim Palmer is a Royal Society Research Professor in Climate Physics at the University of Oxford in the UK.



We think we know why the weather can be so difficult to predict. It's the so-called butterfly effect: The flap of a butterfly's wings in Brazil can set off a tornado in Texas a week later. But because we can't observe all the butterflies in Brazil, we can't reliably predict tornadoes in Texas a week in advance.

As described in James Gleick's masterful 1987 exposition of chaos theory,¹ the discovery of the butterfly effect is generally attributed to MIT meteorologist Edward Lorenz. In 1963 he famously constructed a model of chaos based on three deterministic coupled nonlinear differential equations.² Being chaotic, the evolution of the state of that system is extremely sensitive to the specification of the initial conditions. Therefore, Lorenz's three-component model describes both the butterfly effect and the unpredictability of the weather.

At least, that's the folklore. But it isn't quite correct. The butterfly effect was first described by Lorenz in his talk at the 1972 meeting of the American Association for the Advancement of Science.³ The title was

indeed "Predictability: Does the Flap of a Butterfly's Wings in Brazil Set Off a Tornado in Texas?" In the talk, Lorenz noted that errors in forecasting the position and intensity of low-pressure cyclonic weather systems tend to double every three days or so. Errors in the individual clouds that are embedded in those weather systems, however, tend to double on shorter time scales. And errors in individual eddies in the subcloud turbulence double on time scales shorter still.

The nonlinear Navier-Stokes equations of fluid mechanics couple the subcloud, cloud, and cyclone scales together. Hence, Lorenz noted, even if you could perfectly observe the atmosphere on the 1000 km scale of the low-pressure system, you



THE REAL BUTTERFLY EFFECT

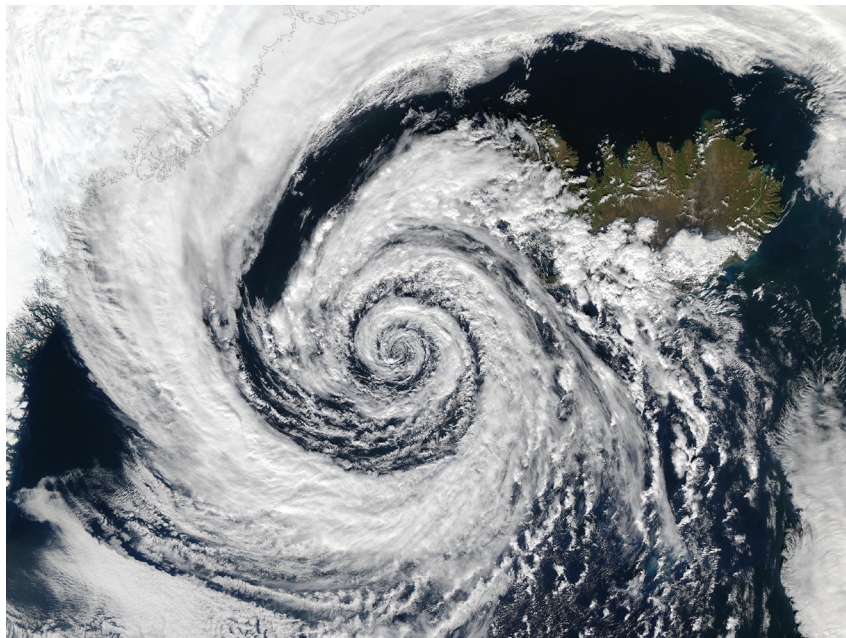


FIGURE 1. A LOW-PRESSURE CYCLONE system contains many individual clouds. Each individual cloud is a turbulent system comprising many small eddies. The real butterfly effect illustrates how uncertainties in the starting conditions for any of those whirls affect our ability to predict the cyclonic system itself. (Courtesy of Jacques Descloitres, MODIS Rapid Response Team, NASA/GSFC.)

would still not be able to predict the structure and intensity of the weather system indefinitely into the future. Initial uncertainties on kilometer or smaller length scales would eventually limit your ability to predict the larger cyclone. The question Lorenz posed was this: How long does it take for uncertainties in the initial conditions on subcloud scales to affect a forecaster's ability to predict position and intensity on the much larger cyclonic scales? (See figure 1.)

The real butterfly effect

Lorenz's 1963 paper cannot address that question—and hence the notion of the butterfly effect as Lorenz intended it to mean in his 1972 talk—because the 1963 model equations do not describe how fluid flows at different spatial scales interact. In fact, in his 1972 talk, Lorenz was informally discussing results from a highly technical paper he had published in 1969 in the Swedish journal *Tellus*. The abstract of the paper, titled “The predictability of a flow which possesses many scales of motion,” begins as follows:

It is proposed that certain formally deterministic fluid systems which possess many scales of motion are observationally indistinguishable from indeterministic systems; specifically, that two states of the system differing initially by a small “observational error” will evolve into two states differing as greatly as randomly chosen states of the system within a finite time interval, which cannot be lengthened by reducing the amplitude of the initial error.⁴

The last clause of the sentence is worth reading a couple of times, because it is so surprising. Lorenz is describing chaotic unpredictability in the extreme. That type of unpredictability is much greater than that in his 1963 model of chaos. In the early model, you can predict as far ahead as you like by making the initial error sufficiently small. From a mathematical standpoint, Lorenz's 1963 model has the property that the evolved

state depends continuously on the initial state. As the initial state tends to the true state, so, too, does the forecast state.

On the basis of the Navier–Stokes partial differential equations, Lorenz's 1969 paper describes systems that do not plausibly have that continuity property. Indeed, the limit of vanishing initial error, which I'll discuss in more detail below, is what's known as a singular limit.

Lorenz, in focus

To better appreciate what Lorenz proposed in his 1969 paper, suppose that we can observe the initial state of the atmosphere perfectly, with no errors or gaps. That does not mean that we can forecast perfectly, because to make a forecast of the weather, you must assimilate observations into a computational weather model, thus creating a set of initial conditions for the model.

The weather model approximates the Navier–Stokes and other relevant atmospheric equations using a finite, 3D array of so-called gridboxes. Collectively, the gridboxes cover the whole atmosphere and oceans. (Some models use finite sets of orthogonal functions, such as spherical harmonics, but that doesn't change the argument.) Inside a gridbox, the weather model erroneously assumes that the atmosphere is completely homogeneous. The horizontal size of each gridbox in the very best global weather-forecast models is currently around 10 km.

Next, let's suppose that we can make accurate weather forecasts of low-pressure systems on average up to seven days ahead with our weather model. In the idealized case of perfect observations, the source of error that limits the forecast's accuracy lies in the gridbox-homogeneity assumption. Hence, it is reasonable to ask (our employers) for a bigger computer that would allow the weather equations to be integrated with a gridbox half the size. The incorrect homogeneity assumption would then be restricted to scales smaller than before by a factor of two.

Would that factor of two double the range of forecast accuracy from 7 days to 14 days? In his 1969 paper, Lorenz argues that it does not. The errors associated with small scales that were unresolved in the old model but are subsequently resolved in the new one would grow faster than errors in the smallest scales resolved in the old model. For example, if the error-doubling time of the newly resolved scales was half the error-doubling time of the previously resolved scales—meaning that the errors grow twice as fast—the predictability time with the new weather model will only increase by a factor of $(1 + \frac{1}{2})$, which is significantly less than a factor of two.

Indeed, if later still we could afford a computer that would



FIGURE 2. PREDICTABILITY in a nonlinear system, such as this Lorenz attractor, is dependent on the initial conditions, whose uncertainties are represented by the size and location of a circular ring. **(a)** The ring of uncertainty does not grow in time at all. **(b)** Started from a lower position, the ring distorts into banana and boomerang shapes, making it unclear whether the actual system undergoes a transition from the left-hand lobe to the right-hand one. **(c)** With the ring initiated almost midway between the lobes, the time evolution of the attractor is now very uncertain, and there is no predictability. (Adopted from ref. 11.)

allow a further halving of the size of the gridboxes, the predictability time would only be increased from $(1 + \frac{1}{2}) \times 7$ days to $(1 + \frac{1}{2} + \frac{1}{4}) \times 7$ days. If you carried on like that—halving the gridbox an infinite number of times—the predictability time would not be infinite. Rather, it would be $(1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} \dots) \times 7$, or 14 days. With infinitesimally small gridboxes, forecasters would have increased the predictability time of the original model by only a factor of two. (The existence of that finite limit is consistent with the Kolmogorov energy spectrum for 3D fluid turbulence.)

Singular limits

But that sounds contradictory. After an infinite number of gridbox halvings, the (now infinitely powerful) computer represents the Navier–Stokes equations precisely. And because those equations are completely deterministic, we should be able to forecast infinitely far ahead.

To understand what accounts for the short forecast range, imagine having a bucket of apples that contain maggots. If you bite into an apple and discover half a maggot, then you have eaten half a maggot—an unpleasant experience. However, if you bite into an apple and discover a quarter of a maggot, then that’s even worse because you have eaten three-quarters of a maggot. More generally, if you bite into an apple and discover $1/n$ of a maggot, you have eaten $1 - 1/n$ of a maggot.

The larger the value of n , the greater the fraction of the maggot you have eaten, and the more unpleasant the experience. You might therefore imagine that the limit $n = \infty$ of a sequence of such apple bitings describes the most unpleasant experience. But it doesn’t. If you bite into an apple and discover no maggot, you may not have eaten a maggot at all! (A tiny maggot fraction is qualitatively different from no maggot.)

That example, first described by theoretical physicist Michael Berry, is known as a singular limit (see his Reference Frame, *PHYSICS TODAY*, May 2002, page 10). Such limits abound in physics. For example, blackbody radiators never experience a UV catastrophe—the prediction that the intensity of their emitted radiation goes to infinity as wavelength decreases—provided

that Planck’s constant h remains nonzero (no matter how small it is). Set h precisely to zero, however, and the classical Rayleigh–Jeans spectrum diverges.

In another example, as long as a fluid’s viscosity remains nonzero, it is able to generate aerodynamic lift across an airfoil, no matter how small the viscosity may be. If viscosity is set to zero, however, the boundary condition across the airfoil qualitatively changes. The lifting force of a 3D body in incompressible, inviscid, irrotational flow is zero, a phenomenon known as d’Alembert’s paradox.

There is also a singular limit at the heart of what I call the real butterfly effect.⁵ No matter how small the initial uncertainty, the butterfly effect limits predictability to a finite time horizon. Only when the initial uncertainty is identically zero can you potentially predict arbitrarily far ahead with the Navier–Stokes equations. That’s an unrealistic limit, of course. Is the singular predictability limit a rigorous mathematical property of the Navier–Stokes equations? No one knows. The problem of whether solutions depend continuously on initial conditions is related to the unsolved Clay Mathematics Institute Millennium Prize Problem concerning the existence of smooth, unique solutions to the Navier–Stokes equations.

Indeterministic results

That is not to say that Lorenz’s more famous 1963 model of chaos has nothing useful to say about the predictability of weather. I have used the model on many occasions to demonstrate that the predictability of a nonlinear system is not a fixed quantity. It varies from one initial condition to another, as shown in figure 2. Hence, although the average predictability of day-to-day weather may be around two weeks, it can sometimes be longer and sometimes shorter than that. Meteorologists can estimate such flow-dependent predictability by running ensembles of forecasts—typically 50 are run from almost but not quite identical initial conditions. When the atmosphere is in a predictable state, the ensemble forecast spread will be relatively small. When the atmosphere is in an unpredictable state, the spread will be relatively large.

THE REAL BUTTERFLY EFFECT

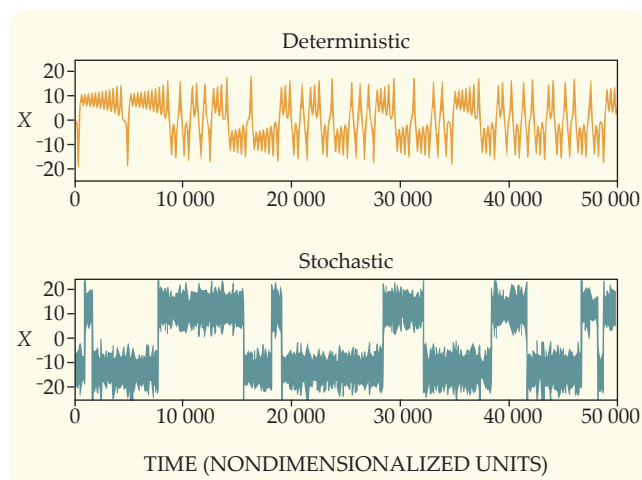


FIGURE 3. ADDING NOISE to Edward Lorenz's 1963 system of equations describing chaos affects its dynamics in a nonintuitive way. The top plot shows a time series of the X variable in the standard (deterministic) Lorenz model. The bottom plot has a much more pronounced structure because noise is present. The noise effectively stabilizes the regimes of the Lorenz attractor, shown in figure 2. (Adapted from ref. 11.)

Ensemble prediction has transformed weather forecasting over recent years. For example, it determines the probability of precipitation on your weather app. More importantly, it is changing the way in which humanitarian and disaster relief agencies respond to extreme weather events. In the past, the unreliability of deterministic predictions meant that they would typically wait for an extreme event to occur before sending in medicine, food, water, and emergency shelter to stricken regions. Now, on the basis of a cost-benefit analysis, those agencies predetermine a threshold probability for extreme weather. And if the ensemble-based forecast probabilities exceed the threshold, the agencies take what's known as "anticipatory action," sending in emergency supplies ahead of the weather event.

The real butterfly effect implies that although the governing partial differential equations are deterministic, any computational representation of the equations will be indeterministic. That's not, however, the way weather and climate models have traditionally been formulated. The processes in such models that cannot be resolved explicitly—cloud formation, the flow of air over small mountains, and ocean mixing, for example—have been represented by deterministic parameterization formulas that mimic molecular viscosity and diffusion.

The real butterfly effect, however, implies that no consistent way to represent those subgrid processes by deterministic formulas exists. One way to alleviate the problem is to make the parameterization formulas in weather and climate models explicitly stochastic.⁶⁷ The first stochastic-parameterization scheme was introduced into a weather forecast model in 1999. And today, most weather models incorporate some form of stochastic parameterization.

Even so, many climate models—even those contributing to assessment reports from the Intergovernmental Panel on Climate Change—are still formulated with deterministic closure schemes. Such models are inconsistent with the Navier–Stokes equations' scaling symmetries, which contributes to their (sometimes substantial) long-term systematic errors.⁸ Stochastic-

ity can have unexpected effects in nonlinear models.⁹ Figure 3, for example, shows that adding noise to the Lorenz 1963 equations helps to stabilize the Lorenz-attractor regimes. The stabilizing effect is quite counterintuitive until you realize that the model makes transitions from one regime to the other in small regions of state space. Those transitions can be disrupted (and thus the regimes stabilized) by small amounts of noise.

Weather forecasting with artificial intelligence?

Artificial intelligence (AI) is now being used to make weather forecasts with levels of skill comparable to more traditional physics-based models. For both training and forecasting, those AI-based models still use sets of gridded, global atmospheric states, in which atmospheric observations have been assimilated into a global physics-based model. Can such AI forecast systems simulate the real butterfly effect?

To answer that question, Tobias Selz and George Craig (both at the German Aerospace Center in Oberpfaffenhofen) compared the growth of estimates of forecast uncertainty using AI and physics-based models last year.¹⁰ The estimate of the initial uncertainty was obtained by taking the difference between two randomly chosen members of an ensemble of data assimilations, which are used in ensemble weather forecasting. The members of the ensemble differ only in the precise values of the observations being assimilated into the model—the variations in those precise values being consistent with observational error.

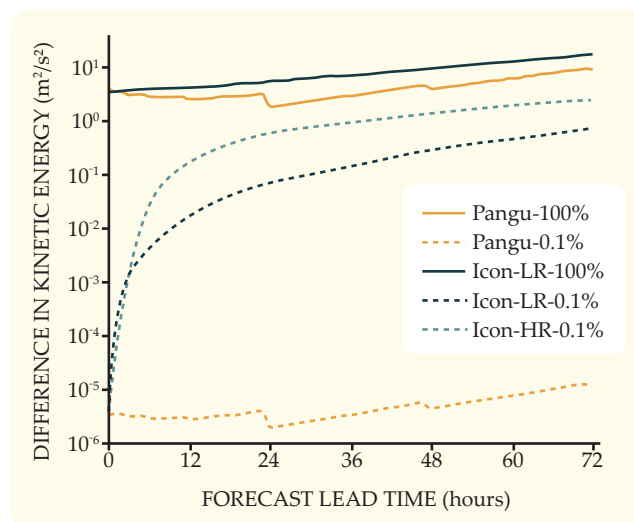


FIGURE 4. THE DIFFERENCE IN A MEASURE of atmospheric kinetic energy between pairs of forecasts as a function of forecast time. The solid black and orange lines show results from a physics-based (Icon) and artificial intelligence (Pangu) model, respectively, when the initial difference between the pairs is comparable with the typical uncertainty in the initial conditions. The dashed lines show differences in kinetic energy when the initial difference is reduced by a factor of 1000. The blue and black dashed lines show the difference in a high- and low-resolution physics-based model, respectively. The orange dashed line shows the lack of growth from an AI model with similar reduced initial perturbation. AI-forecast systems don't capture the physics of the real butterfly effect. (Adapted from ref. 10.)

By construction, the initial error for a weather forecast is spread across a range of scales—from weather systems with a horizontal wavelength of thousands of kilometers down to the model's grid scale of 10 kilometers or so. The theory of data assimilation predicts that if the spacing between atmospheric observations is typically a few tens of kilometers, then observations do well at determining the large-scale initial weather patterns, with little error. On kilometer scales, however, the errors will become almost as large as it is possible for them to be. Small-scale errors in the initial conditions are thus almost saturated, while large-scale errors have plenty of opportunity to grow. Accordingly, errors grow almost immediately at the large scale but not at all at the small scale.

To study the real butterfly effect, Selz and Craig divided the initial-error field by a factor of 1000. Then, the small-scale errors were far from saturated. Because they grow so much faster than the large-scale errors, the errors should be dominated by the small scales. That is precisely what is seen when a physics-based model is used. And Selz and Craig used both a low-resolution and a high-resolution physics-based model to demonstrate it. Figure 4 shows the divergence of pairs of forecasts with small initial differences.

The high-resolution model did a much better job at simulating the rapid growth of the small-scale errors, but the low-resolution model was not completely hopeless; the growth was simply less dramatic. By contrast, the AI system completely failed to predict the growth of small-scale errors. That's perhaps not surprising. In the real world, as I mentioned, the

small-scale errors are already saturated at the initial time. The AI system never learns about the real butterfly effect from its training data. The results demonstrate that you must be cautious when applying AI to the weather-forecast problem; it does not contain the physics of the real butterfly effect.

As I discuss in my book *The Primacy of Doubt*,¹¹ studying the predictability of weather and climate reveals some deep and important properties of nonlinear systems. They are relevant to many problems in applied and fundamental science—in various fields, including social science and the foundations of quantum physics. In short, taking a rigorous approach to the science of uncertainty can help us improve our ability to both predict and understand our very chaotic world.

REFERENCES

1. J. Gleick, *Chaos: Making a New Science*, Viking (1987).
2. E. N. Lorenz, *J. Atmos. Sci.* **20**, 130 (1963).
3. E. N. Lorenz, *The Essence of Chaos*, U. Washington Press (1995), app. 1.
4. E. N. Lorenz, *Tellus* **21**, 289 (1969).
5. T. N. Palmer, A. Döring, G. Seregin, *Nonlinearity* **27**, R123 (2014).
6. T. N. Palmer, *Nat. Rev. Phys.* **1**, 463 (2019).
7. D. Bandak et al., *Phys. Rev. Lett.* **132**, 104002 (2024).
8. T. Palmer, B. Stevens, *Proc. Natl. Acad. Sci. USA* **116**, 24390 (2019).
9. F. Kwasniok, *Philos. Trans. R. Soc. A* **372**, 20130286 (2014).
10. T. Selz, G. C. Craig, *Geophys. Res. Lett.* **50**, e2023GL105747 (2023).
11. T. Palmer, *The Primacy of Doubt: From Climate Change to Quantum Physics, How the Science of Uncertainty Can Help Predict and Understand Our Chaotic World*, Oxford U. Press (2022). PT

PHYSICS TODAY

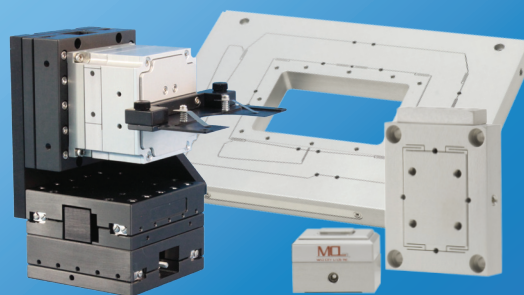
Physics Today Webinars

Encounter A Wide Variety of
Engaging Topics on Leading
Research

Watch Now at
physicstoday.org/webinars

MCL
MAD CITY LABS INC.

Think Nano®
Positioning | Microscopy | Solutions



Closed Loop Nanopositioning
Micropositioning
AFM & NSOM
Single Molecule Microscopes
Custom Solutions

sales@madcitylabs.com • www.madcitylabs.com