Forwards or Backwards

https://forwardorbackwards.com



Can AI Out-Forecast the Weather?

Watch on YouTube: https://www.youtube.com/watch?v=WzDkffC8vXE

For seventy years, a weather forecast has been a physics problem: observations fed into equations, solved on the biggest computer available, grid box by grid box. The European Centre for Medium-Range Weather Forecasts has run that physics every day since 1 August 1979.

That is now changing. Google DeepMind's GraphCast was trained on decades of past weather instead of solving a single equation, and in a December 2023 study in Science it beat ECMWF's own physics-based forecast on ninety per cent of 1,380 separate measures of accuracy. A specialised successor, WeatherNext Cyclones, gave forecasters more than a full extra day of warning on tropical storms, and fed into the US National Hurricane Center's guidance for Hurricane Melissa's rapid intensification ahead of its Category 5 landfall in Jamaica in October 2025 - a result that won Google DeepMind a place in the 2026 Gizmodo Science Fair.

Governments noticed. The US National Weather Service switched on AI-based forecasting systems, fine-tuned from GraphCast, on 17 December 2025, using about three tenths of one per cent of the computing power the physics-based system needs. Environment and Climate Change Canada announced its own hybrid model that April and switched it on that May.

But two 2026 studies - from the Karlsruhe Institute of Technology and Rice University - found the same models still underestimate exactly the events that matter most: record-breaking heat, cold, and the physical structure of the strongest storms, because they have only ever learned from weather that has already happened.

This video traces both halves of that story, and asks how much of forecasting has actually shifted from physics to pattern-matching, and how much still has not.

Every figure is on screen with its source and date.

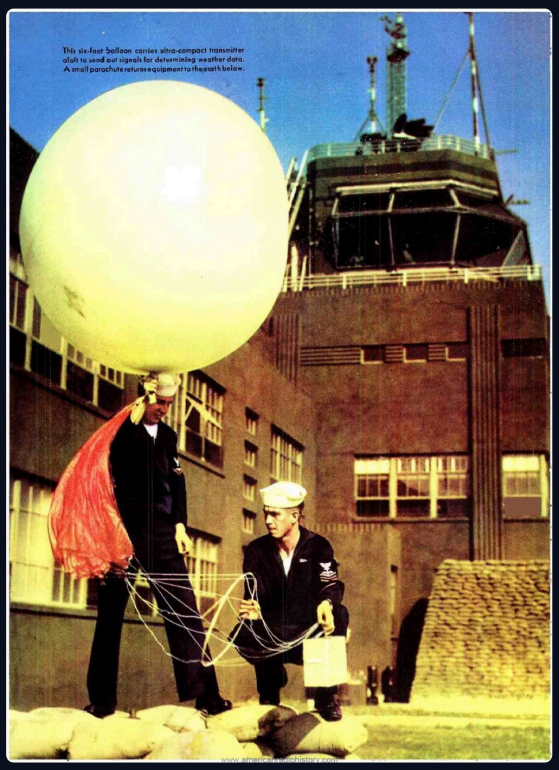
Educational documentary. Not financial or investment advice.

Chapters

1. Seventy years of physics
2. Inside a physics-based forecast
3. The cost of brute force
4. Teaching a model the atmosphere
5. Pattern-matching instead of simulating
6. The head-to-head result
7. AI learns to track hurricanes
8. Tested on a real storm
9. Governments go hybrid
10. Where AI still gets it wrong
11. The hybrid future
12. How much has actually shifted

1. Seventy years of physics





For most of the history of weather forecasting, a forecast was a physics problem, solved by the biggest computer available.

This is a US Navy weather balloon, launched in 1943 to carry instruments up through the atmosphere and radio back readings of pressure, temperature and humidity. Observations like this one are still the raw material every forecast starts from, physics-based or not.

In 1950, the mathematician John von Neumann ran the first computer-based weather forecast on ENIAC, one of the earliest electronic computers. It took the machine about twenty-four hours to calculate a one-day forecast – barely faster than the weather itself. But it proved the idea: feed observations of the atmosphere into equations, and let a machine do the arithmetic.

A quarter of a century later, a group of European governments founded the European Centre for Medium-Range Weather Forecasts, or ECMWF. It was built in November 1975, to share computing power no single country could afford alone. No single national weather service of that era could justify buying a machine big enough on its own, so pooling the cost was the only way to run a forecast this ambitious at all. It did not even own a computer of its own at first; it rented time on someone else's machine until it built its own headquarters near Reading, England. ECMWF has produced an operational forecast, without a gap, every day since the first of August 1979.

That is the starting point for everything that follows: physics equations, solved on a shared machine, running continuously for nearly fifty years. Shipping, aviation and farming all came to depend on it, which is exactly why the next seventy years of this story matter so much.

2. Inside a physics-based forecast





So what actually happens inside one of those forecasts, today?

It begins with observations gathered from everywhere at once — weather balloons, ships, aircraft, and satellites in orbit. A process called data assimilation blends all of those readings with the previous

forecast to build the best possible picture of the atmosphere at this exact moment, over the whole planet.

Observations never cover the whole planet evenly — there are gaps over oceans and deserts where almost nothing is measured directly — so assimilation's job is to fill those gaps with the physically most plausible atmosphere, not just the nearest reading.

That picture is then divided into a three-dimensional grid, a kind of invisible graph paper wrapped around the Earth, with a box for every patch of ground and every layer of the atmosphere above it.

The physics — equations describing how air pressure, temperature and moisture change over time — is solved separately inside every one of those boxes. Then the boxes are made to agree with their neighbours, over and over, stepping forward in time. This is the back of a real supercomputer rack, the kind of machine that does that solving: a forecast like this does not run on a laptop, and it can take hours of computing time to produce a single ten-day forecast.

3. The cost of brute force





The finer the grid, the better the forecast — and that improvement does not come cheap.

This is a Cray-1, a supercomputer from the kind of era ECMWF first relied on. Halving the size of every grid box roughly doubles the detail across the map in two directions, and also forces the

model to take smaller, more frequent time-steps to stay accurate. The computing cost rises far faster than the improvement in detail.

And there is a second limit that no amount of computing power removes. The atmosphere is chaotic. A tiny, unmeasured difference in today's conditions grows into a large difference within a couple of weeks. That is why physics-based forecasts become unreliable past about ten to fourteen days, however fine the grid.

So physics forecasting hit two walls at once: compute that grows faster than detail, and a theoretical ceiling that no amount of compute can buy its way past. Those two walls are exactly where a completely different approach found room to compete.

4. Teaching a model the atmosphere



A different approach set those two walls aside, and asked a different question: instead of solving the equations again every single time, could a machine learn the pattern from everything that had already happened?

Models like Google DeepMind's GraphCast, and the WeatherNext family that followed it, are trained on ERA5. ERA5 is a reanalysis archive that combines decades of past observations into one consistent, gridded record of the atmosphere, going back to 1979.

Nothing in that training solves a physics equation. The model is simply shown an enormous number of past atmospheric states, each one followed by what actually happened next, until it can predict the next state on its own.

That is a genuinely different kind of knowledge from a physics model's. A physics model knows the rules of the air; a trained model knows, from decades of evidence, what tends to happen next.

5. Pattern-matching instead of simulating



The architecture underneath that learning is called a graph neural network – a structure built from points connected to their neighbours, standing in for the atmosphere instead of the rigid grid a physics model solves equations on.

Once trained, the model does not simulate the atmosphere at all. It takes today's conditions as an input and produces tomorrow's as an output, directly – recognising a pattern rather than solving an equation. A forecast that takes a physics model hours of supercomputer time takes one of these trained models well under a minute, on a single processor.

That speed is the whole trade. A physics model earns its forecast by solving equations fresh every time; a trained model earns its forecast once, during training, and then simply recognises the pattern it has already seen. The question the rest of this video has to answer is whether that recognised pattern is actually as good as the equations.

6. The head-to-head result



So does pattern-matching actually forecast as well as solving the physics?

In December 2023, DeepMind and collaborators published a study in the journal Science, comparing GraphCast against ECMWF's own physics-based forecast, called HRES, across 1,380 separate measures of accuracy. GraphCast came out ahead on ninety per cent of them.

That is not a small result. It was the first time a machine-learned model had beaten the world's leading physics-based forecast on the majority of the ways forecasts are actually scored, rather than on a cherry-picked handful.

Before this study, plenty of researchers assumed a model trained on the past would always trail a model that solved the atmosphere's actual equations. This is the paper that made that assumption much harder to defend.

And it did it after training on the past, with no equations solved at the moment the forecast was made — the comparison the previous chapter promised, settled in a single published study.

7. AI learns to track hurricanes





The head-to-head result was for forecasting the atmosphere broadly. Tracking a single hurricane is a narrower, higher-stakes problem, and Google DeepMind built a separate model for it.

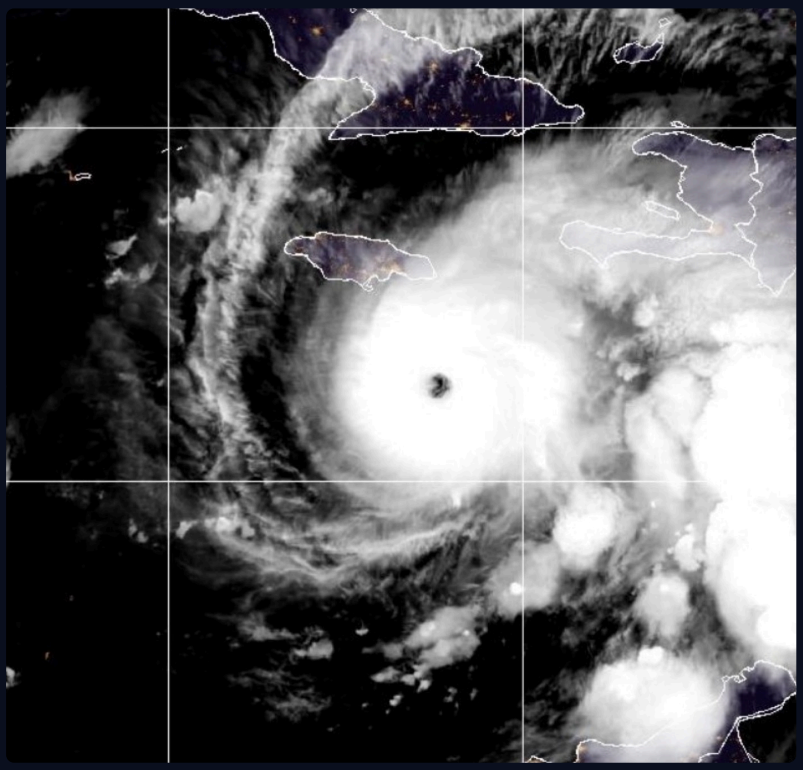
A general forecast can be slightly wrong almost everywhere and still be useful. A hurricane forecast that is slightly wrong about where the storm turns can mean an evacuation order for the wrong town.

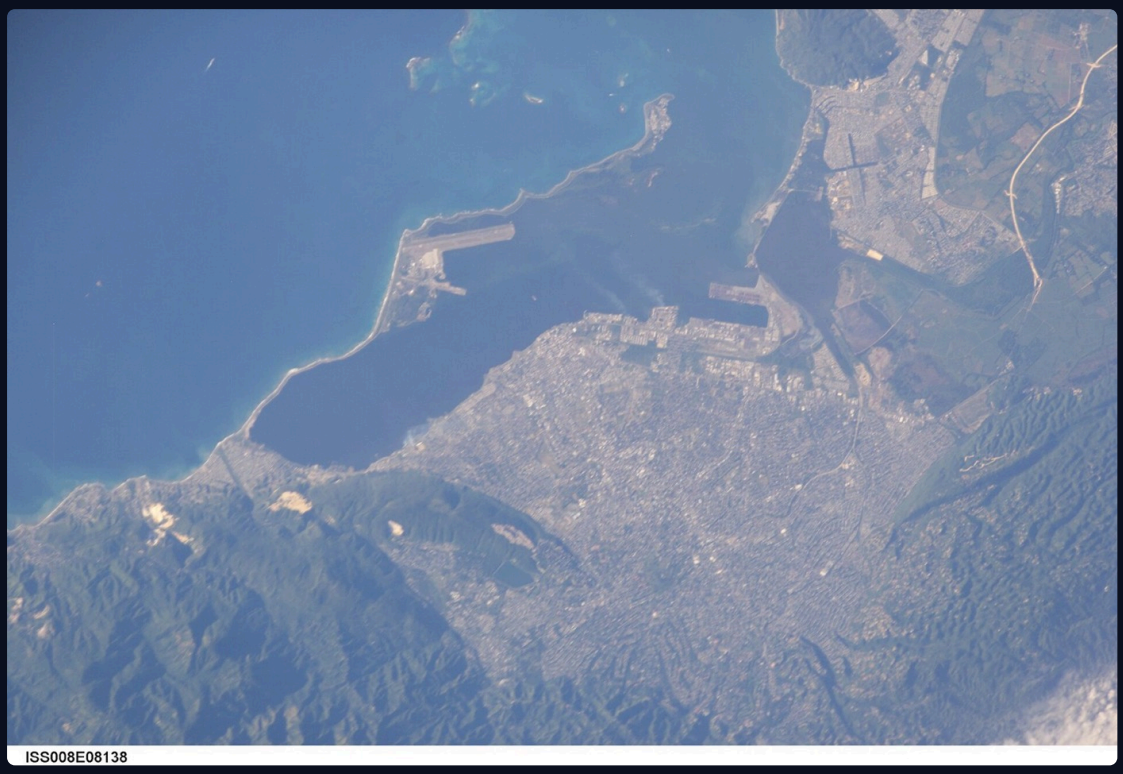
This is the US National Hurricane Center's original building, used from the 1960s into the 1970s. It is the kind of institution that has issued hurricane warnings, by whatever method was available, for decades. WeatherNext Cyclones is the newest attempt to improve what a building like this one can tell the public.

WeatherNext Cyclones was trained on nearly twenty terabytes of atmospheric data, and on IBTrACS, a database of close to five thousand historical storms. Where the general-purpose model looks at the whole atmosphere, this one has only ever studied cyclones, and specialising it that way is what bought the extra accuracy. Google DeepMind reports that its three-day forecasts of a cyclone's track, intensity and size are now about as accurate as two-day forecasts used to be. That is more than a full day of extra warning, on average.

8. Tested on a real storm







This is Hurricane Melissa, photographed from space on 26 October 2025. Days later it struck Jamaica as a Category 5 storm — one of the most powerful Atlantic landfalls ever recorded.

"Our model helped the NHC to make a historic forecast for Hurricane Melissa by predicting the storm's rapid intensification and landfall in Jamaica," Google DeepMind wrote afterward.

Forecasts from WeatherNext Cyclones had been feeding into the Hurricane Center's guidance since the previous June, and Melissa was the storm that put that guidance to its hardest test.

This is Jamaica, photographed from the International Space Station. The island sat directly in the storm's path when Melissa made landfall.

A faster, earlier warning for a storm aimed at an island this size is not an abstract improvement. It is the extra hours a family actually has to leave, and extra time is the one thing no amount of money can buy back once a storm has already made landfall.

9. Governments go hybrid



Within months, two national weather services moved AI models from an experiment into daily, operational use.

On 17 December 2025, the US National Weather Service switched on two new systems. AIGFS gives one best-guess forecast; AIGEFS is an ensemble version, running many slightly different forecasts together to show the range of ways the weather could actually go. Both are fine-tuned from GraphCast on NOAA's own data.

A single sixteen-day AIGFS forecast uses only about three tenths of one per cent of the computing power the older, physics-based system needs. It finishes in around forty minutes, not hours. That difference frees up supercomputer time the physics-based forecast can then spend on the runs AI still cannot do as well.

Four months later, Environment and Climate Change Canada announced its own hybrid model, built on the same GraphCast foundations and blended with Canada's existing physics-based forecast. "Our six-day forecast will be as accurate as our five-day forecast is now, when using the traditional model," the department said, when the system switched on that May.

Two governments, on two continents, reached the same conclusion inside half a year of each other: the AI layer was ready for daily use. Neither switched off its physics-based forecast to do it. They added the AI layer alongside it.

10. Where AI still gets it wrong



None of that means AI has solved forecasting. Two research groups spent 2026 checking exactly where it had not.

Researchers at the Karlsruhe Institute of Technology compared AI models, including GraphCast, against ECMWF's physics-based forecast on genuinely record-breaking heat and cold. The AI models underestimated how extreme those events actually were, because they had only ever been trained on weather that had already happened — events like the 2020 Siberian heatwave sat outside anything in their training record. A physics model has no such blind spot. It solves the same equations whether the result they produce has ever happened before or not.

A separate Rice University study, looking specifically at hurricanes, found that AI forecasts of a storm's track were strong, but its forecasts of a storm's inner structure could look realistic while quietly breaking basic rules of atmospheric physics. "These systems are extraordinarily powerful, but they are not self-validating," said Rice researcher Avantika Gori.

A model trained only on the past has no training data for the record it is about to break. Both findings point at the same gap, from two different angles — one about the biggest numbers on the thermometer, the other about the shape of the storm itself.

11. The hybrid future



Which is why almost nobody serious is arguing to replace physics forecasting with AI outright.

NOAA's own hybrid system, HGEFS, blends an AI-generated ensemble with its existing physics-based ensemble, rather than choosing one over the other. Canada's model keeps its physics-based system running underneath the AI layer specifically to hold onto the small-scale detail AI tends to miss.

The pattern across every government system this video has covered is the same: AI supplies the speed, and physics supplies the grounding, and forecasters are combining the two rather than picking a winner. It is not a transition from one to the other. It is a division of labour, and so far it is holding.

12. How much has actually shifted



So: has weather forecasting shifted from physics to pattern-matching, or is AI just a faster layer riding on the same science underneath?

Both are true, for different parts of the job. For the large-scale, day-to-day forecast – the kind that tells most people whether to carry an umbrella – pattern-matching has already overtaken physics simulation on the government's own numbers.

For the extremes that matter most – the record heatwave, the storm nobody has seen before – physics still has the edge. It is grounded in equations rather than precedent, and every hybrid system built so far keeps it in place for exactly that reason.

Seventy years ago, a forecast was arithmetic, done by hand and then by the earliest computers. Today, part of it is pattern recognition, learned from decades of the atmosphere's own history. Both halves are still running, in the same forecast, at the same time.

Ask whether AI has out-forecast the weather, and for most of what a forecast actually has to say, the honest answer is yes. For the one part where being wrong costs the most, the physics is still doing the job nothing has yet replaced it for.

Sources and credits

Primary sources

- Google DeepMind, 'WeatherNext AI model achieves breakthrough in forecasting cyclones,' deepmind.google, 6 Aug 2026, reporting a same-day Nature paper.
- Lam et al., 'Learning skillful medium-range global weather forecasting,' Science, 14 Dec 2023.
- NOAA/NWS Service Change Notice 25-89 (AIGFS, AIGEFS, HGEFS v1.0), implemented 17 Dec 2025; corroborated by HPCwire, 17 Dec 2025.
- Environment and Climate Change Canada, canada.ca press releases, 9 Apr 2026 and 5 May 2026.
- Karlsruhe Institute of Technology press release, via energiezukunft.eu, 5 May 2026.
- Rice University news, 'AI weather models show promise for hurricane forecasts, but new Rice study finds key physical limitations,' 11 Mar 2026, reporting a study in Journal of Geophysical Research: Atmospheres.
- Gizmodo, 2026 Science Fair coverage of Google DeepMind's WeatherNext Cyclones.

Photo credits (Wikimedia Commons)

- US sailors launching a radiosonde weather balloon, 1943: Unknown authorUnknown author, Public domain - https://commons.wikimedia.org/wiki/File%3ALaunching_radiosonde_1943.jpg
- Cray-1 supercomputer, Computer History Museum: NecroBones at English Wikipedia, Public domain - https://commons.wikimedia.org/wiki/File%3ACray-1_%281%29.jpg
- Rear of a server rack, NERSC data center: Derrick Coetzee from Berkeley, CA, USA, CC0 - https://commons.wikimedia.org/wiki/File%3ARear_of_rack_at_NERSC_data_center_-_closeup.jpg
- Hurricane Melissa satellite image, 26 October 2025: Unknown authorUnknown author, Public domain - https://commons.wikimedia.org/wiki/File%3AHurricane_Melissa_2025-10-26.png
- View of Jamaica from the ISS: Earth Science and Remote Sensing Unit, Lyndon B. Johnson Space Center, Public domain - https://commons.wikimedia.org/wiki/File%3AISS008-E-8138_-_View_of_Jamaica.jpg
- Original National Hurricane Center building, 1964-1978: ESSA, progenitor of NOAA, Public domain - https://commons.wikimedia.org/wiki/File%3ANational_Hurricane_Center_%28Old_Building%29_1964.png

Not regulated financial advice.