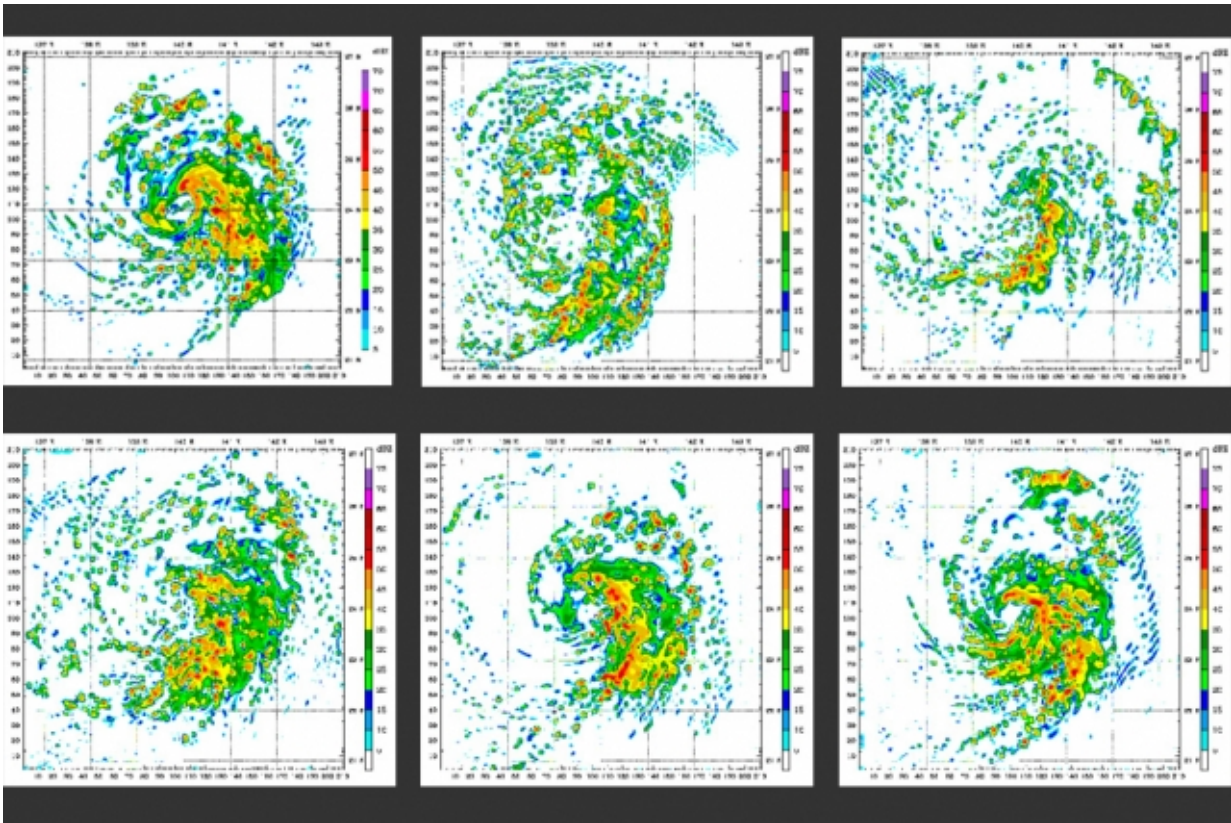# 'Shrinking bull's-eye' algorithm speeds up complex modeling from days to hours

November 16 2015, by Jennifer Chu

Stills from the Weather Research and Forecasting Model. Credit: Wikipedia/Almoz

To work with computational models is to work in a world of unknowns: Models that simulate complex physical processes—from Earth's

changing climate to the performance of hypersonic combustion engines—are staggeringly complex, sometimes incorporating hundreds of parameters, each of which describes a piece of the larger process.

Parameters are often question marks within their models, their contributions to the whole largely unknown. To estimate the value of each unknown parameter requires plugging in hundreds, if not thousands, of values, and running the model each time to narrow in on an accurate value—a computation that can take days, and sometimes weeks.

Now MIT researchers have developed a new algorithm that vastly reduces the computation of virtually any computational model. The algorithm may be thought of as a shrinking bull's-eye that, over several runs of a model, and in combination with some relevant data points, incrementally narrows in on its target: a probability distribution of values for each unknown parameter.

With this method, the researchers were able to arrive at the same answer as a classic computational approaches, but 200 times faster.

Youssef Marzouk, an associate professor of aeronautics and astronautics, says the algorithm is versatile enough to apply to a wide range of computationally intensive problems.

"We're somewhat flexible about the particular application," Marzouk says. "These models exist in a vast array of fields, from engineering and geophysics to subsurface modeling, very often with unknown parameters. We want to treat the model as a black box and say, 'Can we accelerate this process in some way?' That's what our algorithm does."

Marzouk and his colleagues—recent PhD graduate Patrick Conrad, Natesh Pillai from Harvard University, and Aaron Smith from the University of Ottawa—have published their findings this week in the

Journal of the American Statistical Association.

Modeling "Monopoly"

In working with complicated models involving multiple unknown parameters, computer scientists typically employ a technique called Markov chain Monte Carlo (MCMC) analysis—a statistical sampling method that is often explained in the context of the board game "Monopoly."

To plan out a monopoly, you want to know which properties players land on most often—essentially, an unknown parameter. Each space on the board has a probability of being landed on, determined by the rules of the game, the positions of each player, and the roll of two dice. To determine the probability distribution on the board—the range of chances each space has of being landed on—you could roll the die hundreds of times.

If you roll the die enough times, you can get a pretty good idea of where players will most likely land. This, essentially, is how an MCMC analysis works: by running a model over and over, with different inputs, to determine a probability distribution for one unknown parameter. For more complicated models involving multiple unknowns, the same method could take days to weeks to compute an answer.

Shrinking bull's-eye

With their new algorithm, Marzouk and his colleagues aim to significantly speed up the conventional sampling process.

"What our algorithm does is short-circuits this model and puts in an approximate model," Marzouk explains. "It may be orders of magnitude

cheaper to evaluate."

The algorithm can be applied to any complex model to quickly determine the probability distribution, or the most likely values, for an unknown parameter. Like the MCMC analysis, the algorithm runs a given model with various inputs—though sparingly, as this process can be quite time-consuming. To speed the process up, the algorithm also uses relevant data to help narrow in on approximate values for unknown parameters.

In the context of "Monopoly," imagine that the board is essentially a three-dimensional terrain, with each space represented as a peak or valley. The higher a space's peak, the higher the probability that space is a popular landing spot. To figure out the exact contours of the board—the probability distribution—the algorithm rolls the die at each turn and alternates between using the computationally expensive model and the approximation. With each roll of the die, the algorithm refers back to the relevant data and any previous evaluations of the model that have been collected.

At the beginning of the analysis, the algorithm essentially draws large, vague bull's-eyes over the board's entire terrain. After successive runs with either the model or the data, the algorithm's bull's-eyes progressively shrink, zeroing in on the peaks in the terrain—the spaces, or values, that are most likely to represent the unknown parameter.

"Outside the normal"

The group tested the algorithm on two relatively complex models, each with a handful of unknown parameters. On average, the algorithm arrived at the same answer as each model, but 200 times faster.

"What this means in the long run is, things that you thought were not

tractable can now become doable," Marzouk says. "For an intractable problem, if you had two months and a huge computer, you could get some answer, but you would not necessarily know how accurate it was. Now for the first time, we can say that if you run our algorithm, you can guarantee that you'll find the right answer, and you might be able to do it in a day. Previously that guarantee was absent."

Marzouk and his colleagues have applied the algorithm to a complex model for simulating movement of sea ice in Antarctica, involving 24 unknown parameters, and found that the algorithm is 60 times faster arriving at an estimate than current methods. He plans to test the algorithm next on models of combustion systems for supersonic jets.

"This is a super-expensive model for a very futuristic technology," Marzouk says. "There might be hundreds of unknown parameters, because you're operating outside the normal regime. That's exciting to us."

More information: Accelerating Asymptotically Exact MCMC for Computationally Intensive Models via Local Approximations. arxiv.org/abs/1402.1694

This story is republished courtesy of MIT News (web.mit.edu/newsoffice/), a popular site that covers news about MIT research, innovation and teaching.

Provided by Massachusetts Institute of Technology

Citation: 'Shrinking bull's-eye' algorithm speeds up complex modeling from days to hours (2015, November 16) retrieved 2 May 2024 from https://phys.org/news/2015-11-bull-eye-algorithm-complex-days-hours.html