Machine listening for earthquakes: Algorithms pick out hidden signals that could boost geothermal energy production

May 23 2018

Using machine-learning tools to analyze 46,000 earthquakes at The Geysers geothermal field in California, researchers discovered patterns that correspond to water-injection flows, suggesting a link to the mechanical processes that produce earthquakes Credit: Stepheng3

For all that seismologists have learned about earthquakes, new technologies show how much remains to be discovered.

In a new study in *Science Advances*, researchers at Columbia University show that machine learning algorithms could pick out different types of earthquakes from three years of [earthquake] recordings at The Geysers in California, one of the world's oldest and largest [geothermal reservoirs]. The repeating patterns of earthquakes appear to match the seasonal rise and fall of water-injection flows into the hot rocks below, suggesting a link to the mechanical processes that cause rocks to slip or crack, triggering an earthquake.

"It's a totally new way of studying earthquakes," said study coauthor Benjamin Holtzman, a geophysicist at Columbia's Lamont-Doherty Earth Observatory. "These machine learning methods pick out very subtle differences in the raw data that we're just learning to interpret."

The approach is novel in several ways. The researchers assembled a catalog of 46,000 earthquake recordings, each represented as energy waves in a seismogram. They then mapped changes in the waves' frequency through time, which they plotted as a spectrogram—a kind of musical roadmap of the waves' changing pitches, were they to be converted to sound. Seismologists typically analyze seismograms to estimate an earthquake's magnitude and where it originated. But looking at an earthquake's frequency information instead allowed the researchers to apply machine-learning tools that can pick out patterns in music and human speech with minimal human input. With these tools, the researchers reduced each earthquake to a spectral "fingerprint" reflecting its subtle differences from the other quakes, and then used a clustering algorithm to sort the fingerprints into groups.

The machine-learning assist helped researchers make the link to the fluctuating amounts of water injected belowground during the energy-extraction process, giving the researchers a possible explanation for why the computer clustered the signals as it did. "The work now is to examine these clusters with traditional methods and see if we can understand the

physics behind them," said study coauthor Felix Waldhauser, a seismologist at Lamont-Doherty. "Usually you have a hypothesis and test it. Here you're building a hypothesis from a pattern the machine has found."

If the earthquakes in different clusters can be linked to the three mechanisms that typically generate earthquakes in a geothermal reservoir—shear fracture, thermal fracture and hydraulic cracking—it could be possible, the researchers say, to boost power output in geothermal reservoirs. If engineers can understand what's happening in the reservoir in near real-time, they can experiment with controlling water flows to create more small cracks, and thus, heated water to generate steam and eventually electricity. These methods could also help reduce the likelihood of triggering larger earthquakes—at The Geysers, and anywhere else fluid is pumped underground, including at fracking-fluid disposal sites. Finally, the tools could help identify the warning signs of a big one on its way—one of the holy grails of seismology.

The research grew out of an unusual artistic collaboration. As a musician, Holtzman had long been attuned to the strange sounds of earthquakes. With sound designer Jason Candler, Holtzman had converted the seismic waves of recordings of notable earthquakes into sounds, and then speeded them up to make them intelligible to the human ear. Their collaboration, with study coauthor Douglas Repetto, became the basis for Seismodome, a recurring show at the American Museum of Natural History's Hayden Planetarium that puts people inside the earth to experience the living planet.

As the exhibit evolved, Holtzman began to wonder if the human ear might have an intuitive grasp of earthquake physics. In a series of experiments, he and study coauthor Arthur Paté, then a postdoctoral researcher at Lamont-Doherty, confirmed that humans could distinguish between temblors propagating through the seafloor or more rigid

continental crust, and originating from a thrust or strike-slip fault.

Encouraged, and looking to expand the research, Holtzman reached out to study coauthor John Paisley, an electrical engineering professor at Columbia Engineering and Columbia's Data Science Institute. Holtzman wanted to know if machine-learning tools might detect something new in a gigantic dataset of earthquakes. He decided to start with data from The Geysers because of a longstanding interest in geothermal energy.

"It was a typical clustering problem," says Paisley. "But with 46,000 earthquakes it was not a straightforward task."

Paisley came up with a three-step solution. First, a type of topic modeling algorithm picked out the most common frequencies in the dataset. Next, another algorithm identified the most common frequency combinations in each 10-second spectrogram to calculate its unique acoustic fingerprint. Finally, a clustering algorithm, without being told how to organize the data, grouped the 46,000 fingerprints by similarity. Number crunching that might have taken a computer cluster several weeks was done in a few hours on a laptop thanks to another tool, stochastic variational inference, Paisley had earlier helped develop.

When the researchers matched the clusters against average monthly water-injection volumes across The Geysers, a pattern jumped out: A high injection rate in winter, as cities send more run-off water to the area, was associated with more earthquakes and one type of signal. A low summertime injection rate corresponded to fewer earthquakes, and a separate signal, with transitional signals in spring and fall.

The researchers plan to next apply these methods to recordings of other naturally occurring earthquakes as well as those simulated in the lab to see if they can link signal types with different faulting processes. Another study published last year in Geophysical Research Letters

suggests they are on a promising track. A team led by Los Alamos researcher Paul Johnson showed that machine learning tools could pick out a subtle acoustic signal in data from laboratory experiments and predict when the next microscopic earthquake would occur. Though natural faults are more complex, the research suggests that machine learning could lead to insights for identifying precursors to big earthquakes.

More information: "Machine learning reveals cyclic changes in seismic source spectra in Geysers geothermal field" *Science Advances* (2018). advances.sciencemag.org/content/4/5/eaao2929

Provided by Columbia University