# Using sparse data to predict lab quakes

December 17 2021



Stick-slip events in the earth cause damage like this, but limited data from these relatively rare earthquakes makes them difficult to model with machine learning. Transfer learning may provide a path to understanding when such deep faults slip. Credit: Dreamstime

A machine-learning approach developed for sparse data reliably predicts fault slip in laboratory earthquakes and could be key to predicting fault slip and potentially earthquakes in the field. The research by a Los

Alamos National Laboratory team builds on their previous success using data-driven approaches that worked for slow-slip events in earth but came up short on large-scale stick-slip faults that generate relatively little data—but big quakes.

"The very long timescale between [major earthquakes](#) limits the [data sets](#), since major faults may slip only once in 50 to 100 years or longer, meaning seismologists have had little opportunity to collect the vast amounts of observational data needed for [machine learning](#)," said Paul Johnson, a geophysicist at Los Alamos and a co-author on a new paper, "Predicting Fault Slip via Transfer Learning," in *Nature Communications*.

To compensate for limited data, Johnson said, the team trained a [convolutional neural network](#) on the output of [numerical simulations](#) of laboratory quakes as well as on a small set of data from lab experiments. Then they were able to predict fault slips in the remaining unseen lab data.

This research was the first application of transfer learning to numerical simulations for predicting fault slip in lab experiments, Johnson said, and no one has applied it to earth observations.

With transfer learning, researchers can generalize from one model to another as a way of overcoming data sparsity. The approach allowed the Laboratory team to build on their earlier data-driven machine learning experiments successfully predicting slip in laboratory quakes and apply it to sparse data from the simulations. Specifically, in this case, transfer learning refers to training the neural network on one type of data—simulation output—and applying it to another—experimental data—with the additional step of training on a small subset of experimental data, as well.

"Our aha moment came when I realized we can take this approach to

earth," Johnson said. "We can simulate a seismogenic fault in earth, then incorporate data from the actual fault during a portion of the slip cycle through the same kind of cross training." The aim would be to predict fault movement in a seismogenic fault such as the San Andreas, where data is limited by infrequent earthquakes.

The team first ran numerical simulations of the lab quakes. These simulations involve building a mathematical grid and plugging in values to simulate fault behavior, which are sometimes just good guesses.

For this paper, the convolutional neural network comprised an encoder that boils down the output of the simulation to its key features, which are encoded in the model's hidden, or latent space, between the encoder and decoder. Those features are the essence of the input data that can predict fault-slip behavior.

The neural network decoded the simplified features to estimate the friction on the fault at any given time. In a further refinement of this method, the model's latent space was additionally trained on a small slice of experimental data. Armed with this "cross-training," the neural network predicted fault-slip events accurately when fed unseen data from a different experiment.

Provided by Los Alamos National Laboratory