SAND2017-8350C

Deep Learning Waveform Source Locations

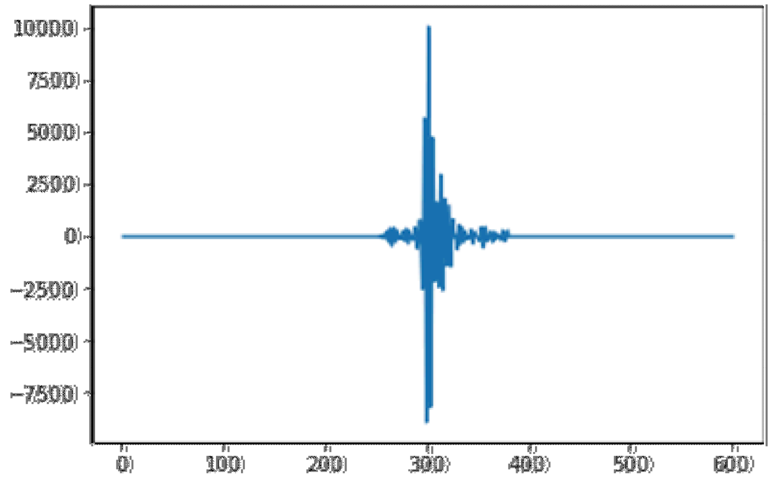
Nick Blazier / 6362

Antonio Gonzales / 6362

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.

Problem

- *Seismic events are constantly happening all around the globe.*
- *These events are detected at seismic sensor stations as waveforms.*
- *The task of detecting an event in the noise and then determining the source location is difficult and usually done by experts by hand.*
- *Automated systems are often used as a good starting estimate to aid analysts*



Goal

Evaluate if Deep Learning can help

- Can Neural Networks be applied to 1D waveforms to return latitude/longitude source positions
- Does it even make sense?
  - Due to the complex structure of the earth events that happen in relatively close proximity may look nothing alike
  - Can we get clean data?

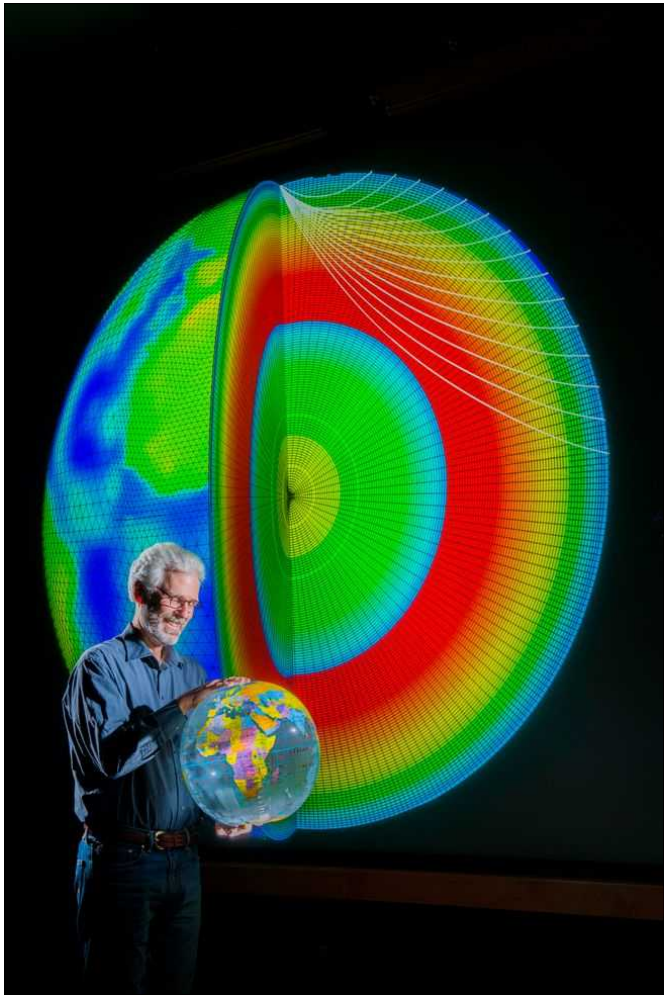


Photo Credit: Sandia National Labs: Randy Montoya

Challenges

- Since seismic events are a natural phenomenon usually occurring near plate boundaries, we don't have samples across the entire globe.
  - Current state of the art solutions look for similar historical events (waveform correlation) or use complex earth models to correlate events seen across networks of sensors.
  - Labelling new events isn't feasible for waveform correlation
    - It is unlikely DL will help in this regard
- The same source waveform will appear different at each station it is detected based on what parts of the earth it travels through
  - We will only train on one stations to see if we can learn source locations. A full system would need models trained individually on EVERY station.

Approach

Framework: Keras

Hardware: DGX

Model: Variant of VGG modified to fit our waveforms

Data: 259,514 Hand labeled seismic waveforms for one seismic stations(courtesy of GNEM)

Input: Waveforms: 380 points, 9 waveforms per sample

Output: latitude and longitude

-unit vectors

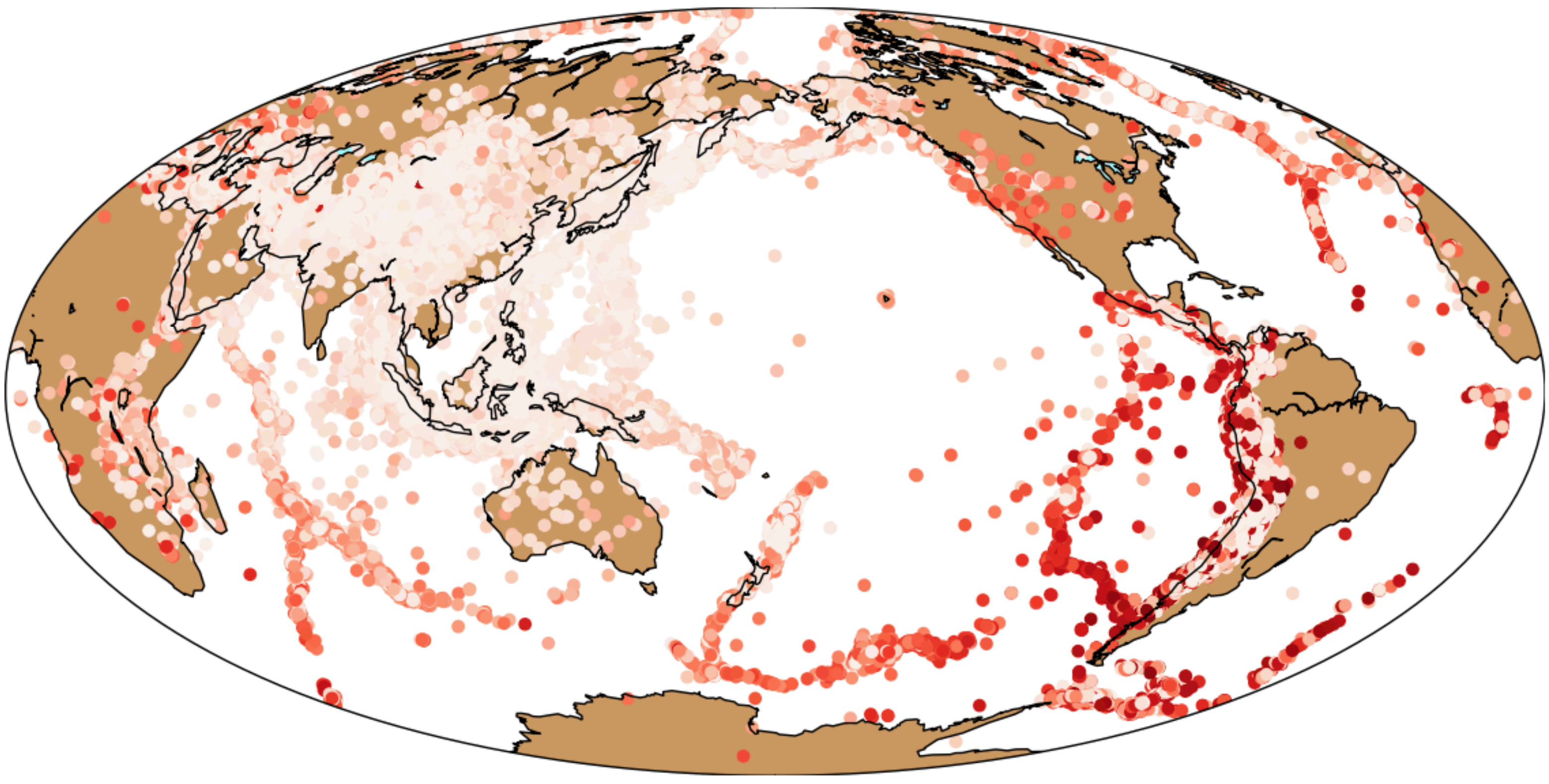
-classifications

Results

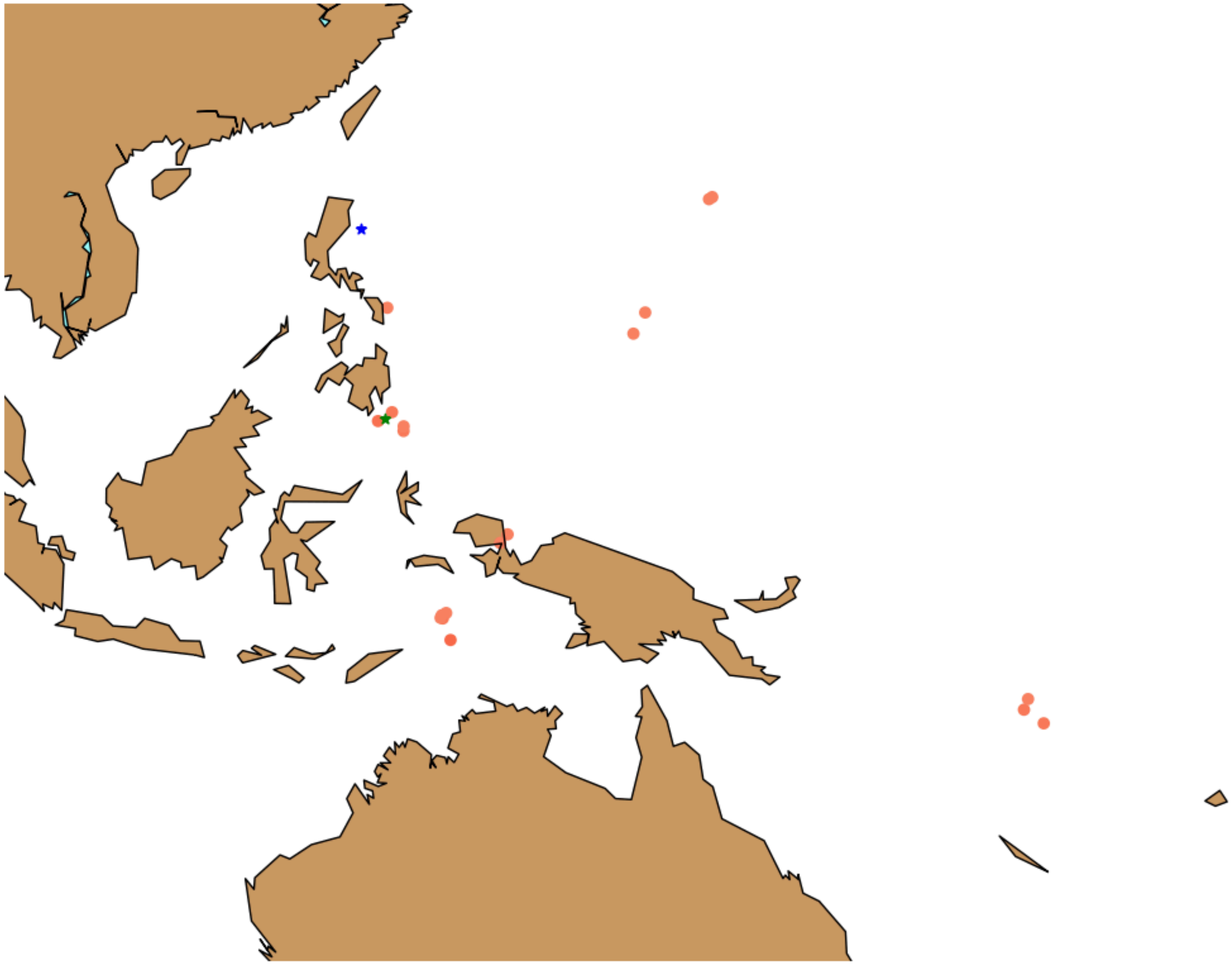
- Still looking over results but initial results are interesting
- VGG like model trained
  - Top 10% within 72km
  - Top 20% within 133km
  - Top 30% within 207km
- Good:
  - Correctly labels many examples ANN struggles with
- Bad:
  - Misses on locations with many nearby examples
- Much to learn
  - Need to look into exactly what it is/isn't good at
  - Insights into how it gets "miracle" examples
  - Can we determine a confidence in the prediction

Note: Correct = < 2 degrees from source (~222km)

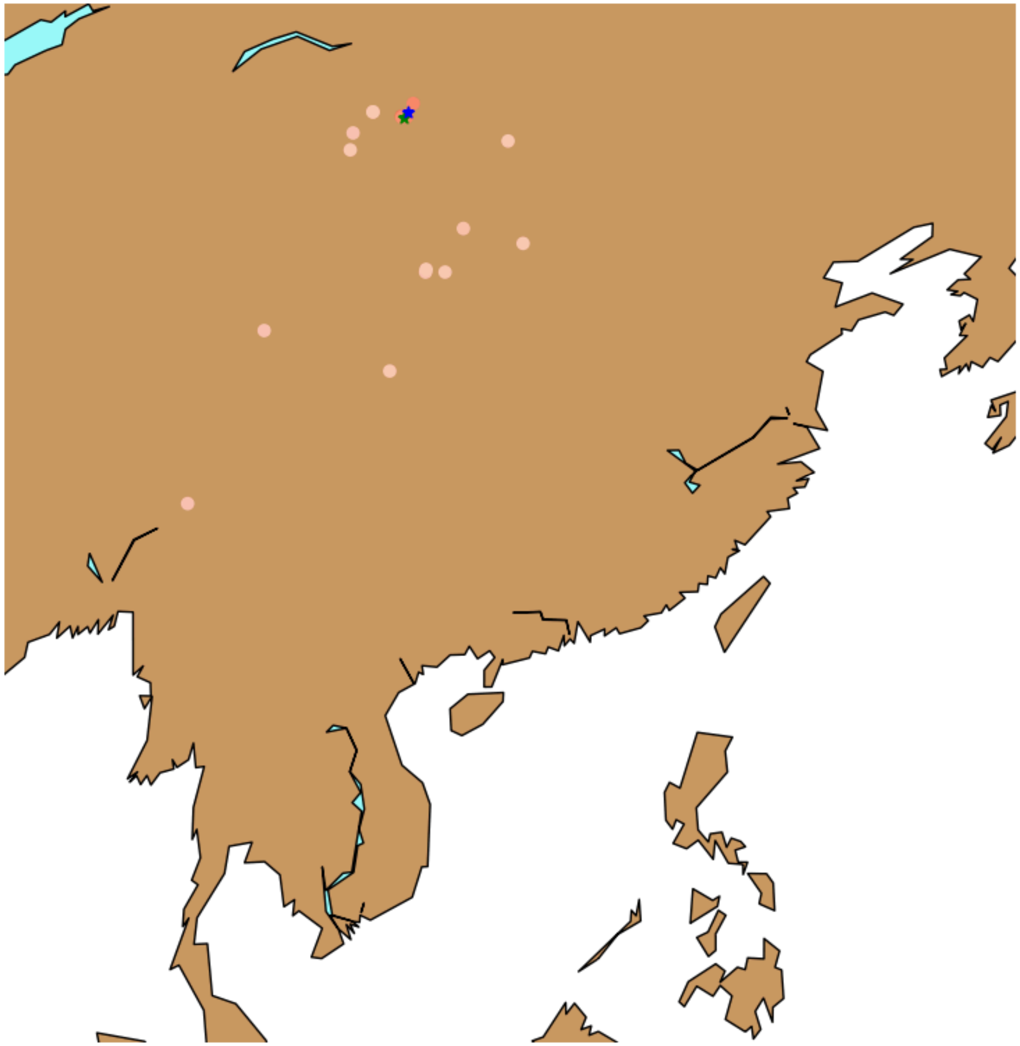
Map of data



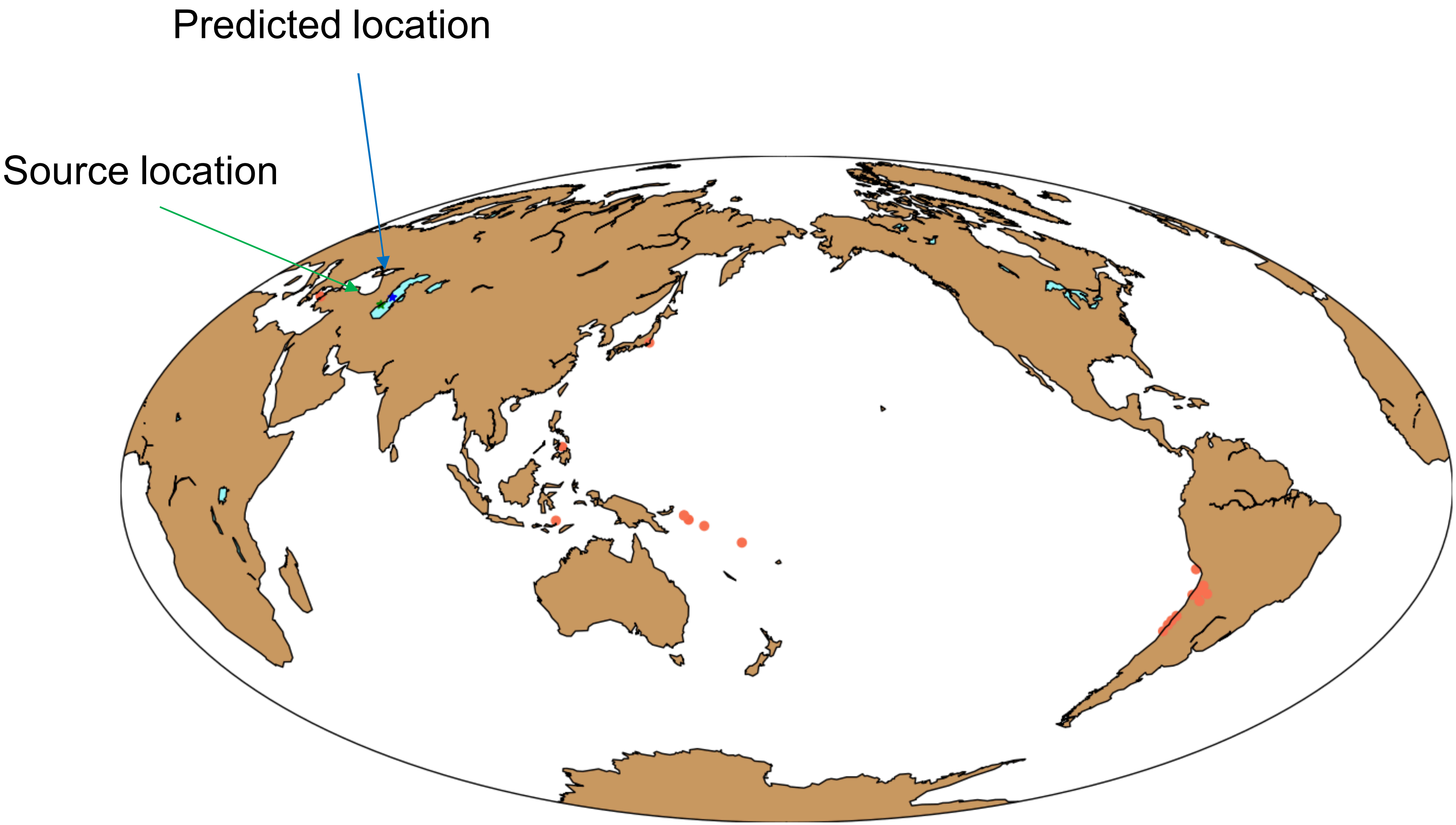
Bad Result



Good Result



Miracle Result



Future work

- Looking for funding to further investigate
  - Modifying current high performing architectures to work on our 1D data
  - Creating our own specialized models
  - Discretizing the globe and “classifying” each event
  - Analysis of events it locates correctly
  - Analysis of events it locates incorrectly
  - Analysis on smaller, “cleaner” stations
  - Try models that incorporate time better
    - LSTM
    - RNNs