T4-06: NET-VISA Model and Inference Improvements

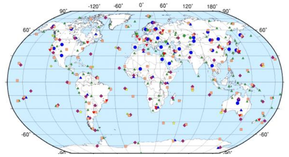
Nimar Arora, Stuart Russell, Paul Kidwell, Erik Sudderth
University of California Berkeley, Lawrence Livermore National Lab, Brown University.

CTBTO PREPARATORY COMMISSION

Introduction

• Global seismic monitoring for the **Comprehensive Nuclear-Test-Ban Treaty (CTBT)** aims to recover the time, location, depth, and magnitude for all seismic events in the magnitude range of interest.

• Data from the International Monitoring System (IMS) are processed in real time at the International Data Centre (IDC) in Vienna



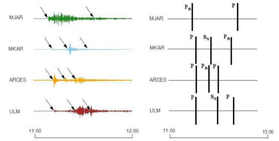
Blue dots and triangles are primary seismic stations.

• Our goal is to improve the sensitivity and accuracy of automated processing at IDC

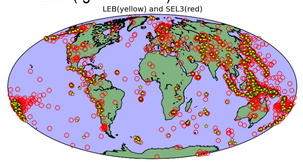
• NET-VISA (NETwork processing by Vertically Integrated Seismic Analysis) has reduced detection failures by more than half, with no reduction in accuracy

The Problem

• IDC records roughly 10000 detections a day of which 90% are spurious, i.e., small local events or detector noise
• Many real events (i.e., magnitude 2 or higher) are not detected at all.
• Data association problem: Which true events caused which observed detections?



• The current automated system (**SEL3**) detects 69% of real events and creates twice as many spurious (nonexistent) events; somewhat unreliable below mag. 3.5
• 16 human analysts find more events, correct existing ones, throw out spurious events, generate **LEB** (“ground truth”)

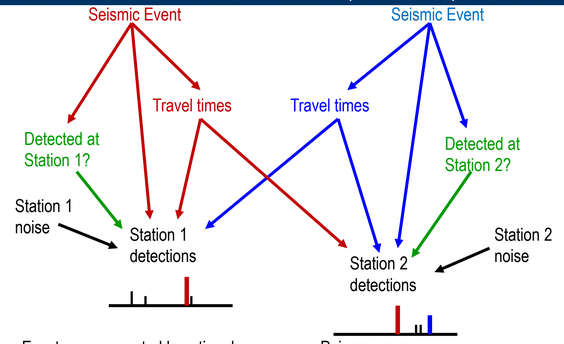


Yellow stars – LEB, Red circles – SEL3. Results for 1 week.

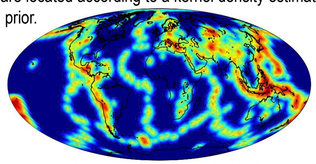
Our Approach

• Unlike SEL3, which processes the data in stages, we propose a single vertically integrated probability model.
• Our model is empirically estimated and includes seismic knowledge as prior information.

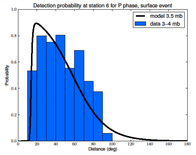
Generative Model of Seismic Event Generation, Transmission, and Detection

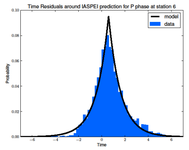


• Events are generated by a time-homogenous Poisson process.
• Earthquakes are located according to a kernel density estimate while explosions have a uniform prior.



• Event magnitudes are distributed according to the Gutenberg-Richter distribution (exponential distribution with rate log(10)).
• Event detection probabilities depend on the station, the seismic wave type (phase), event magnitude, and distance from the event to the station.
• Event parameters – arrival time, azimuth, amplitude, etc. – have station-specific distributions.

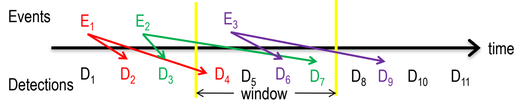




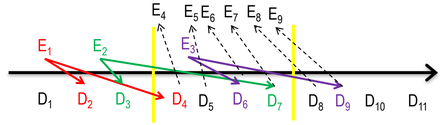
• Noise detections are generated by a station-specific time-homogenous Poisson process.
• All parameters are estimated from historical training data.

Inference

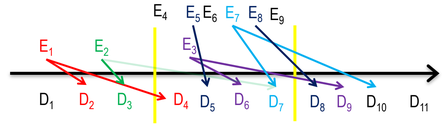
• Given the set of detections at all the stations, we need to infer **the most probable explanation (MPE)** – a sequence of events and the association of events to detections.
• Inference works by modifying the current world through a sequence of moves which mainly focus on events and detections in the current window.



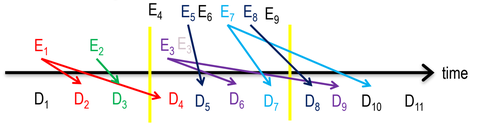
• The birth move adds new events by probabilistically “inverting” detections



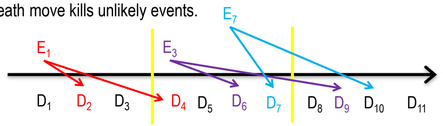
• The re-associate move shuffles detections among the events.



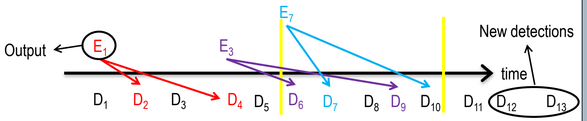
• The relocate move changes event locations.



• The death move kills unlikely events.



• The window moves forward, new detections are added and old events are output

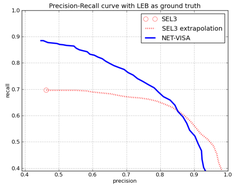


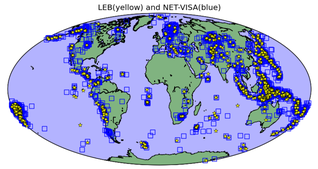
Results

• SEL3: Current automated seismic bulletin.
• LEB: Analyst bulletin starting from SEL3.
• NET-VISA: MPE with generative model.

m_b range	SEL3		NET-VISA	
	Recall	Error (km)	Recall	Error (km)
0 – 2	64.9	101	89.2	106
2 – 3	50.0	186	86.1	140
3 – 4	66.5	104	86.2	121
> 4	86.6	70	93.9	77

• Precision and recall computed using max-cardinality bipartite matching with LEB (assumed to be ground truth).
• Average error is the average distance between matched events.





SEL3 extrapolation is based on scores from an SVM trained on true and false SEL3 events (Mackey, Kleiner, and Jordan . *AGU 2009*)

• Results are based on a 3 month dataset of which 1 week was used for validation.
• LEB is not perfect as the following results suggest.
• In the continental United States of 33 events reported by NEIC:
  • LEB got 4 correct out of 4 predicted events
  • NET-VISA got 7 correct out of 35 predicted events
• In Japan out of 1565 events reported by JMA:
  • LEB got 29 correct out of 29 predictions
  • NET-VISA got 33 correct out of 52 predictions
• In Europe out of 101 events reported by PRU
  • LEB got 5 correct out of 10 predictions
  • NET-VISA got 11 correct out of 43 predictions
• In Central Asia out of 101 events reported by NNC
  • LEB got 35 correct out of 74 predictions
  • NETVISA got 50 correct out of 166 predictions

Source of NEIC, JMA, and PRU, NNC events was the ISC bulletin.

Conclusions

• NET-VISA reduces detection failures by more than a factor of 2 compared to SEL3
• Currently being tested in the CTBTO vDEC environment for possible deployment
• Next step: SIG-VISA extends generative model down to signal level.

The authors were supported by contracts from the Preparatory Commission for the CTBT and the U.S. Department of Energy.