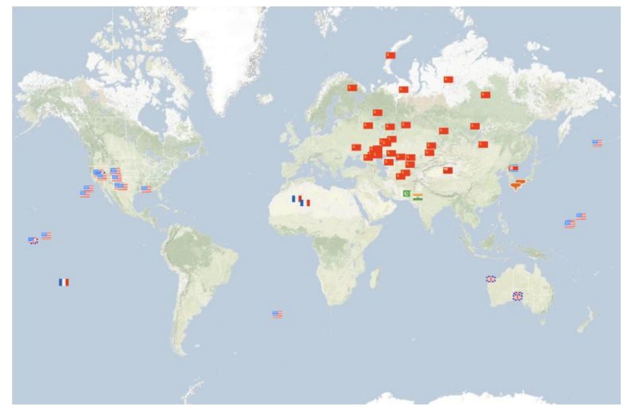
Global seismic monitoring as probabilistic inference

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

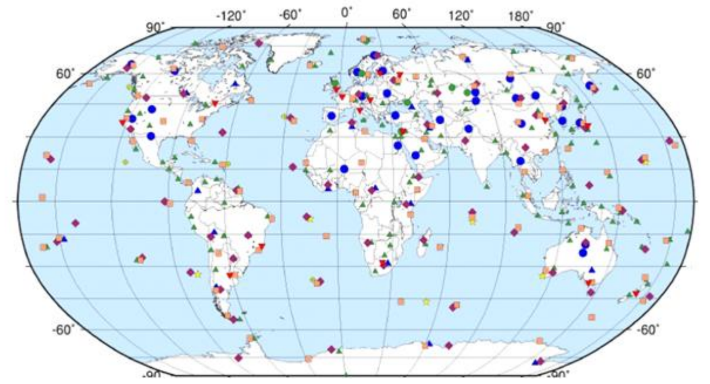
Introduction

• The **Comprehensive Nuclear-Test-Ban Treaty (CTBT)** bans all nuclear explosions on Earth whether for military or for peaceful purposes.



Locations of all known nuclear explosions.

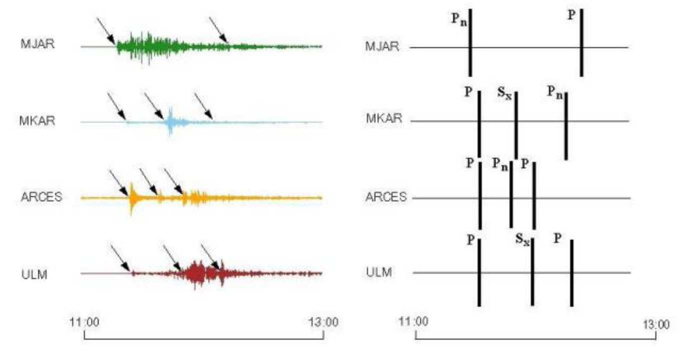
•. A sparse global network of seismic, hydroacoustic, infrasound, and radionuclide stations monitors the earth for potential violations.
• We focus here on seismic events.



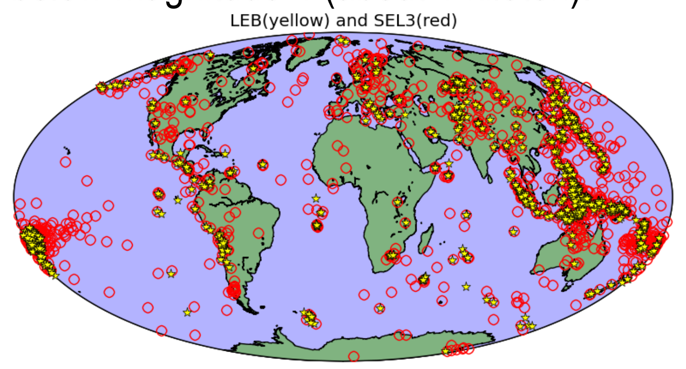
Blue dots and triangles are primary seismic stations.

The Problem

• Roughly 10000 detections a day of which 90% are spurious, i.e. small local events like passing trains, falling trees, etc.
• Many real events (i.e. magnitude 2 or higher) are not detected at all.
• Data Association Problem: What were the true events given the observed detections?



• The current automated system (**SEL3**) detects 69% of real events and creates twice as many spurious (nonexistent) events.
• 16 human analysts find more events, correct existing ones, throw out spurious events, generate **LEB** ("ground truth")
• Unreliable below magnitude 4 (about 1 kiloton).

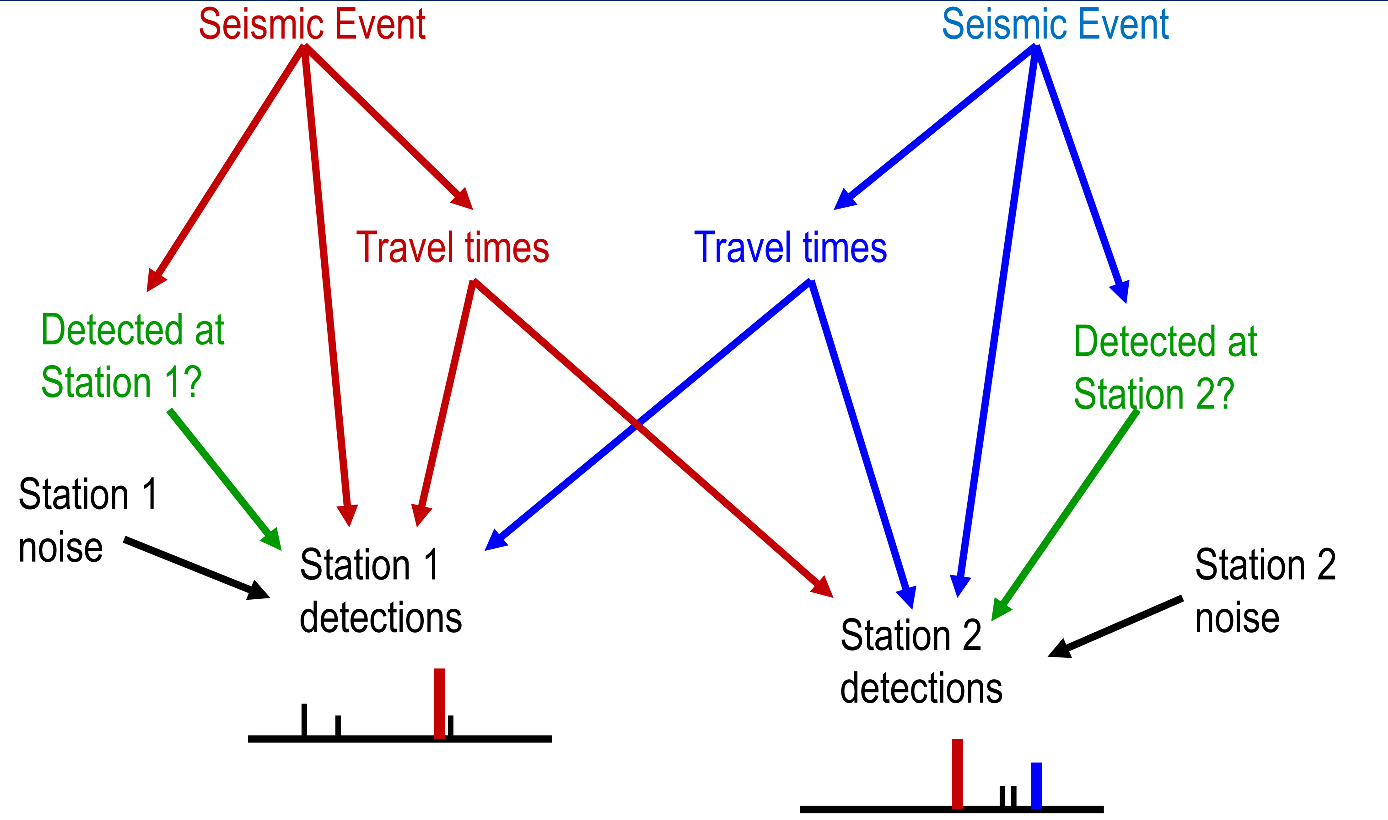


Yellow stars – LEB, Red circles – SEL3. Results for 1 week.

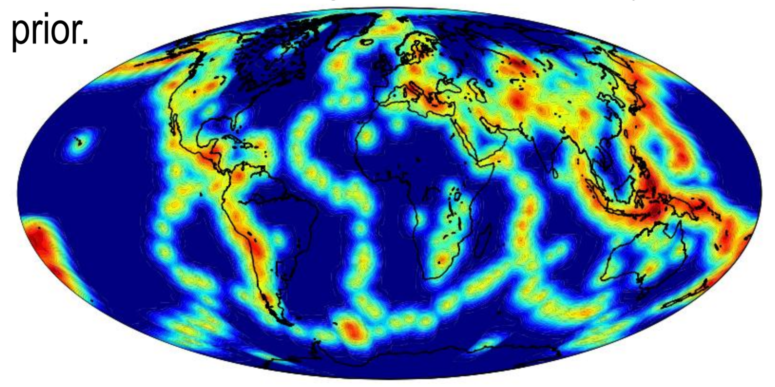
Our Approach

• 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.

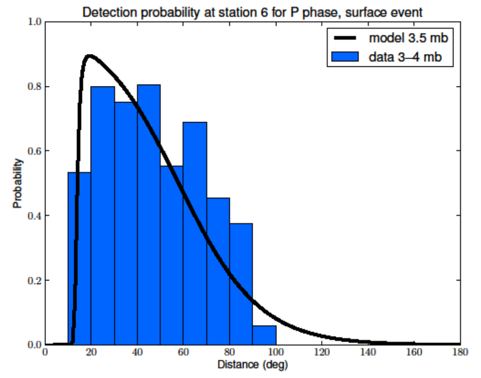
Generative Model of Seismic Event Generation, Transmission, and Detection

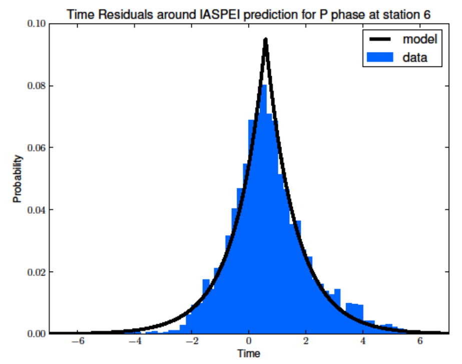


• Events are generated by a time-homogenous Poisson process.
• Earthquakes are located according to a kernel density estimate while explosions have a uniform prior.



• Event magnitudes are distributed as per the Guttenberg 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.

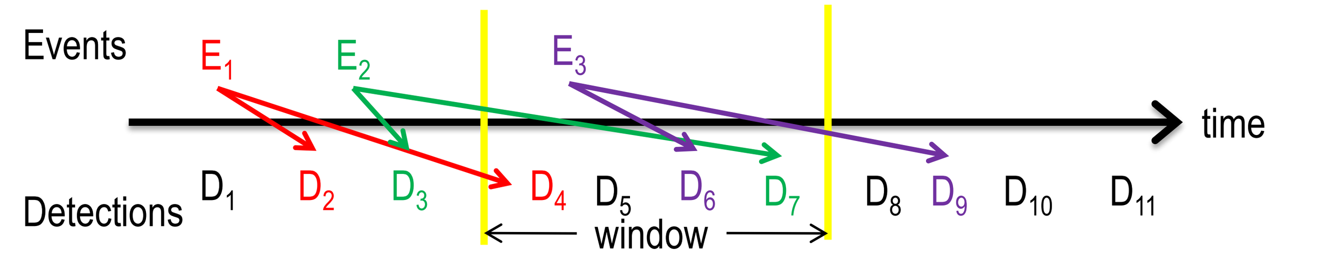




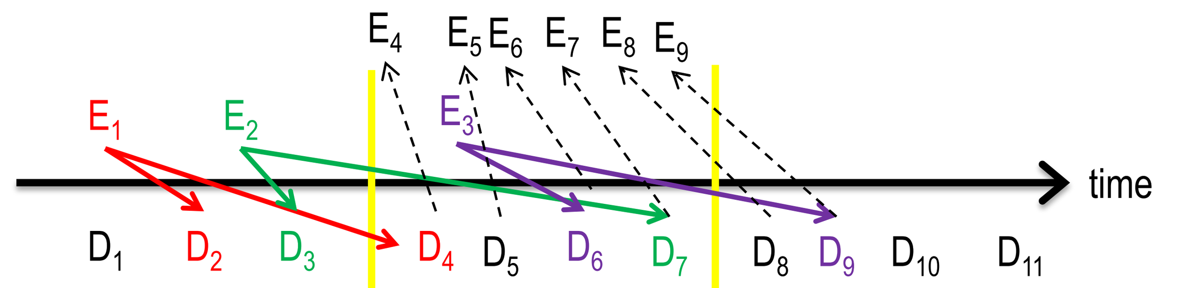
• Noise detections are generated by a station-specific time-homogenous Poisson process.
• All parameters are estimated from historical training data.

Inference

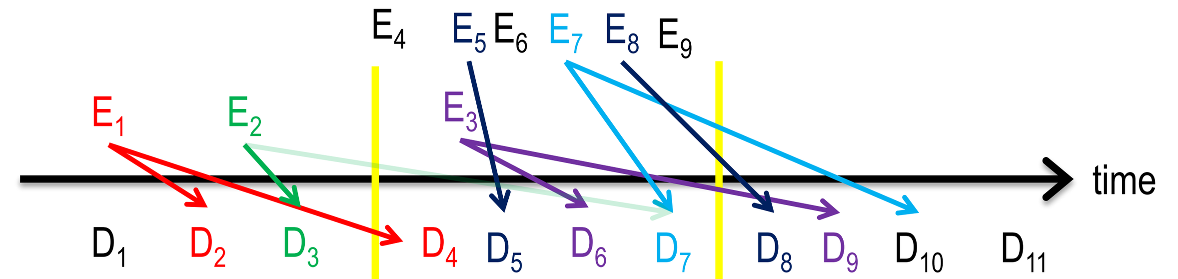
• Given the set of detections at all the stations, we need to infer **the most probable explanation (MPE)** – a sequence of events and the association of events to detections.
• Inference works by modifying the current world through a sequence of moves which mainly focus on events and detections in the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



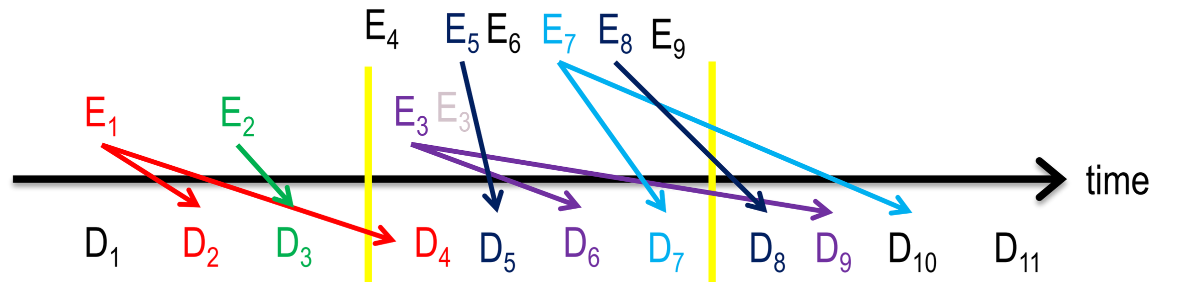
• The birth move adds new events by probabilistically "inverting" detections



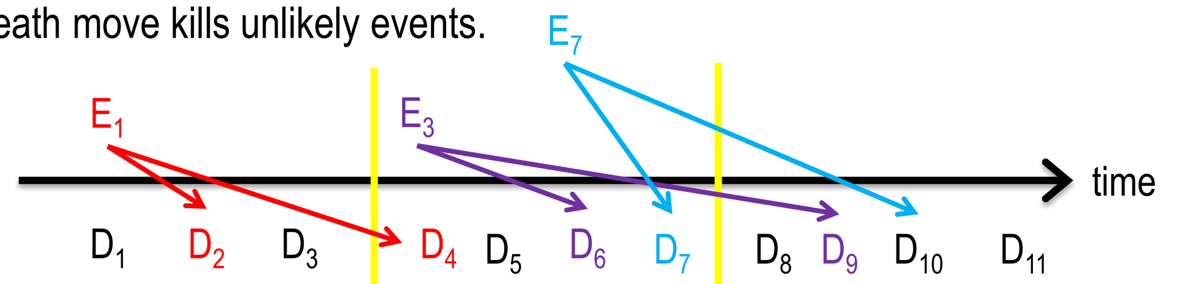
• The re-associate move shuffles detections among the events.



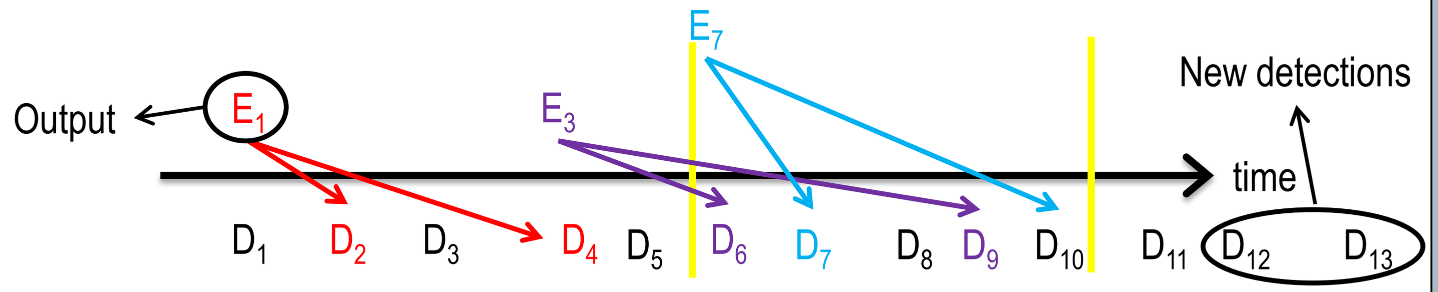
• The relocate move changes event locations.



• The death move kills unlikely events.



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

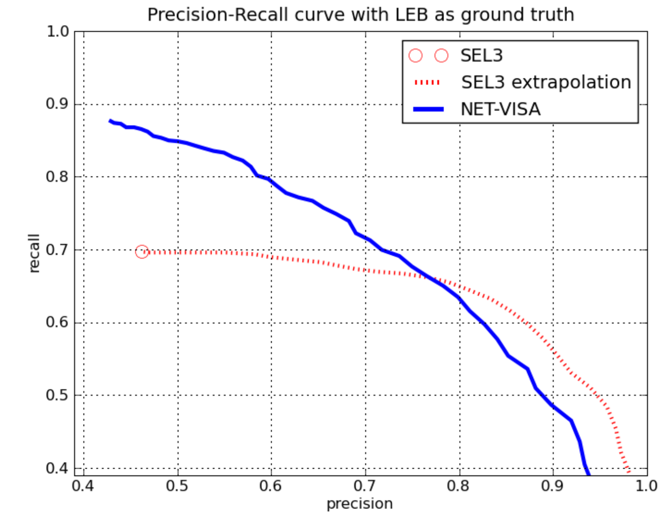


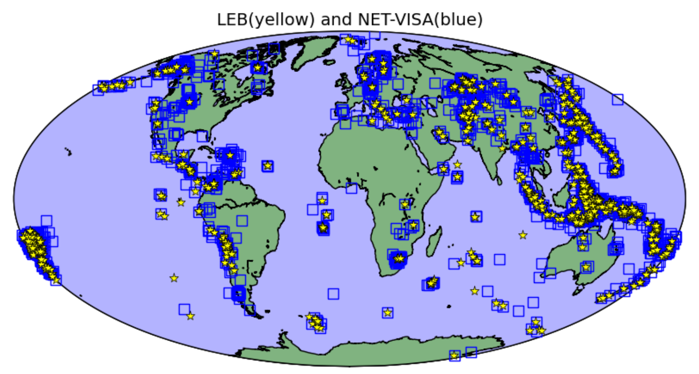
Results

• 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	87
2 – 3	50.0	186	80.6	134
3 – 4	66.5	104	85.8	106
> 4	86.6	70	93.9	70

• 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, the table below shows the performance of LEB and NET-VISA for prediction on the continental United States with the USGS as ground truth.

LEB		NET-VISA	
Precision	Recall	Precision	Recall
4/4	4/33	9/51	9/33

Conclusions

• NET-VISA is twice as effective as SEL3.
• Under consideration for deployment by CTBTO.
• 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.