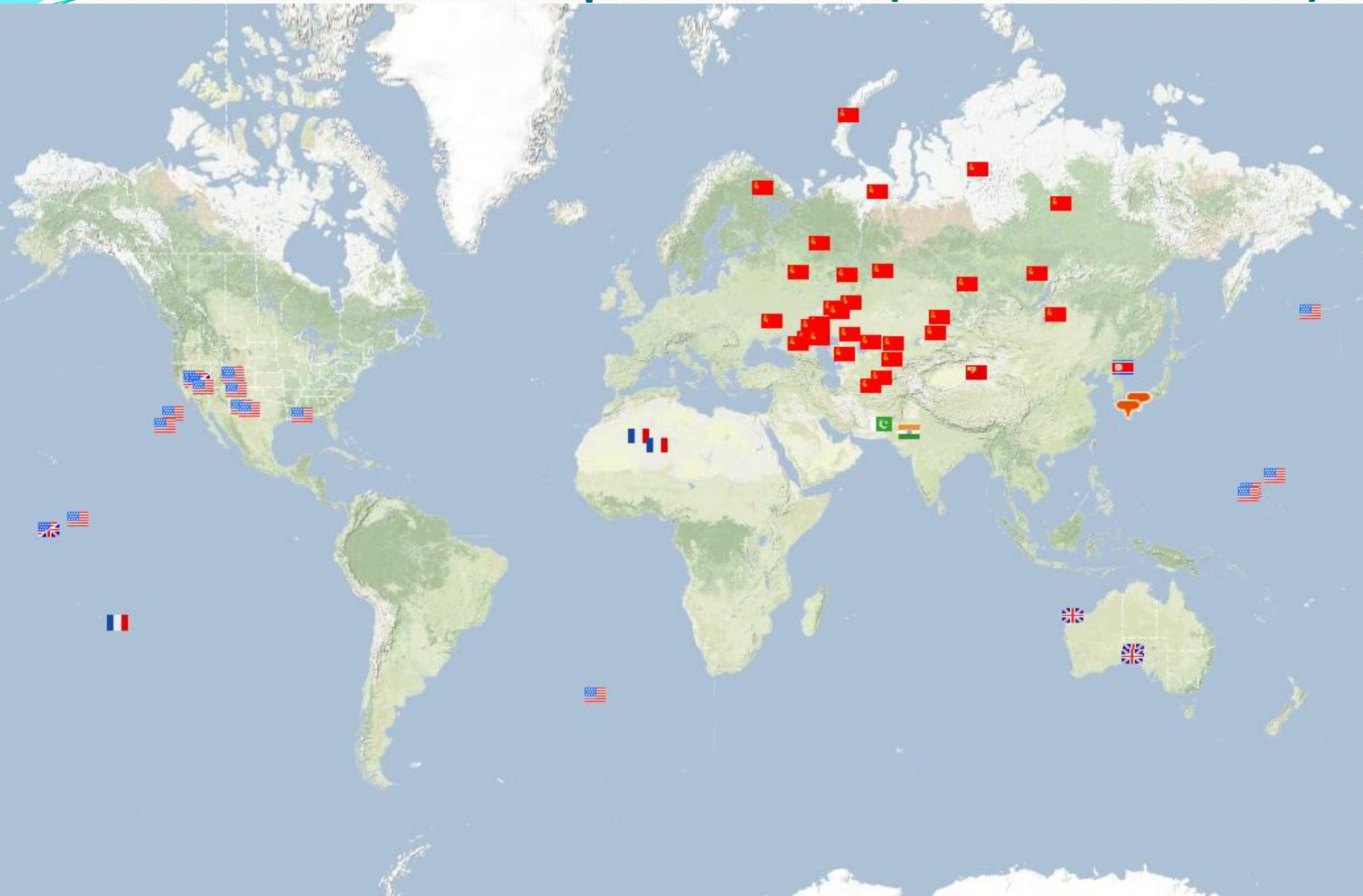
Global seismic monitoring: A Bayesian approach

Nimar Arora

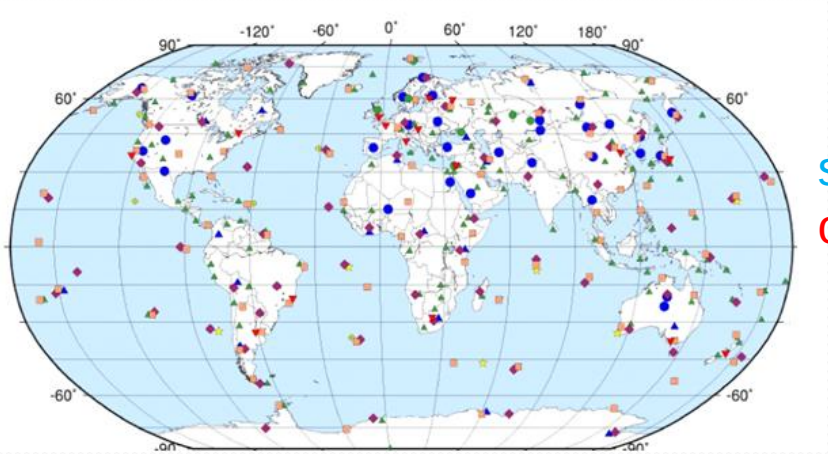
Computer Science, UC Berkeley

Joint work with Stuart Russell, Erik Sudderth, and Paul Kidwell

2053 Nuclear Explosions (1945 - 2009)



CTBT & IMS

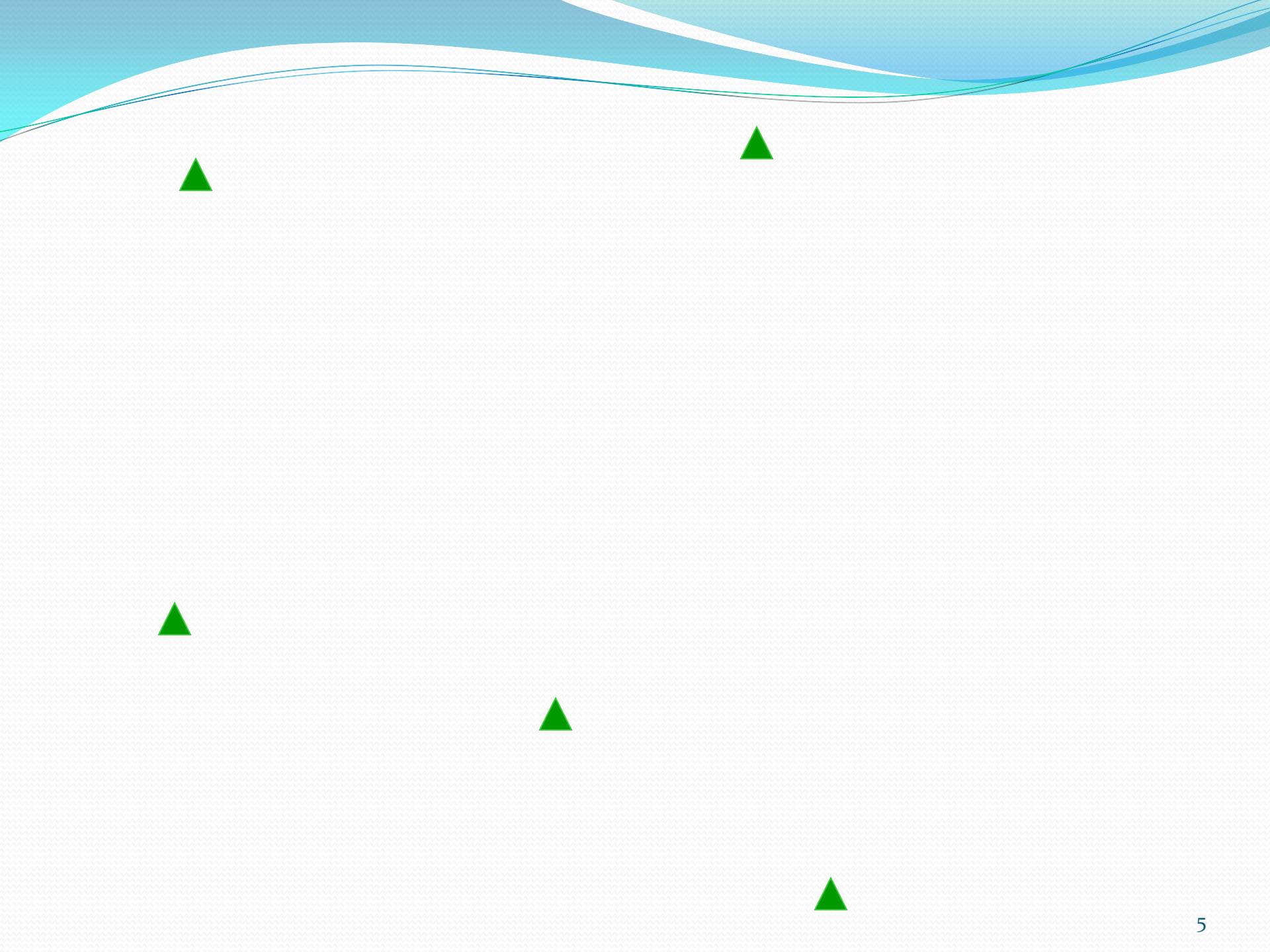


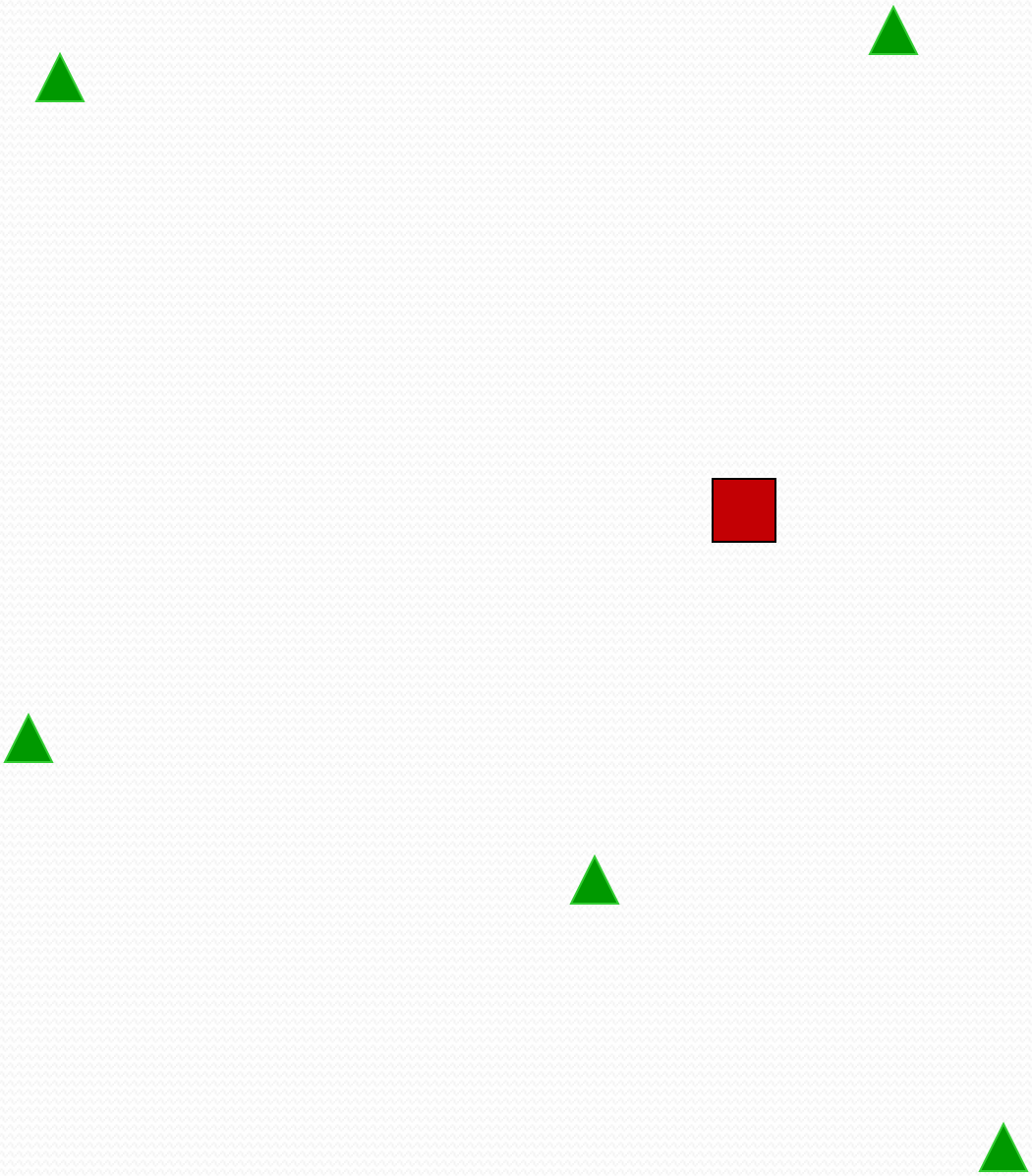
seismic stations
other stations

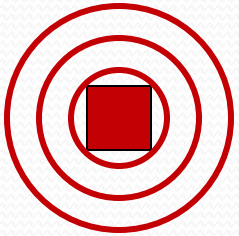
- Bans all testing of nuclear weapons on earth
  - Allows for outside inspection of 1000km²
- ~ 110 seismic stations in IMS (International Monitoring System)
- Need 9 more ratifications including US, China
- US Senate refused to ratify in 1998
  - "too hard to monitor"

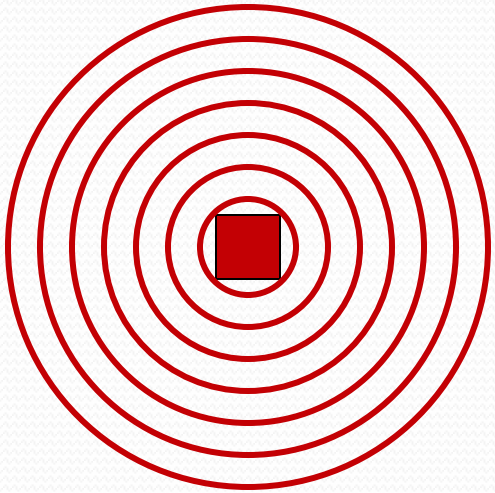
Overview

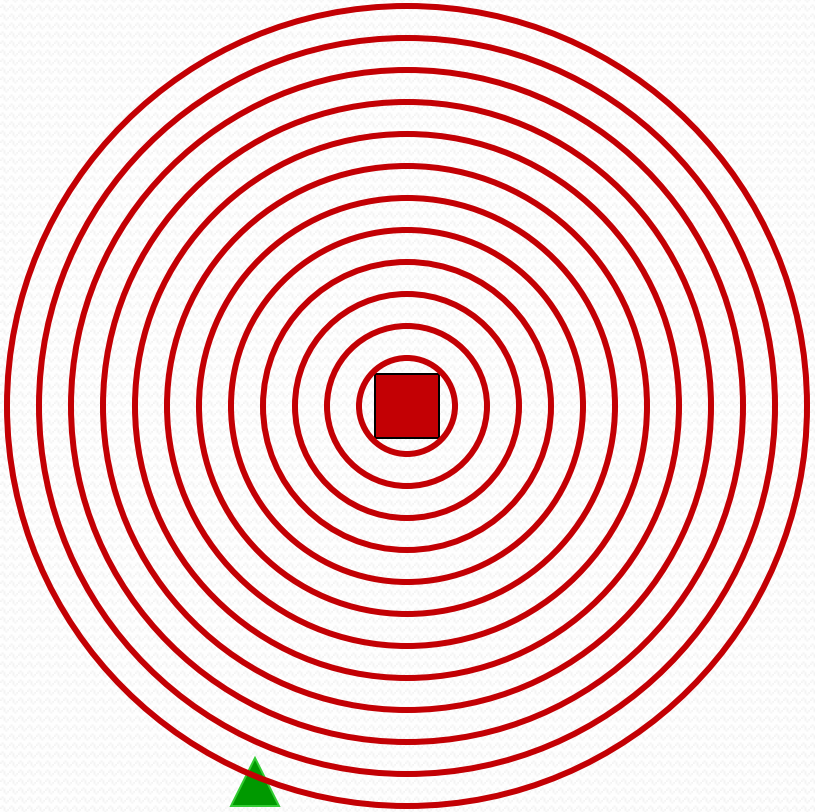
- Seismology
- Generative Model
- Inference
- Results







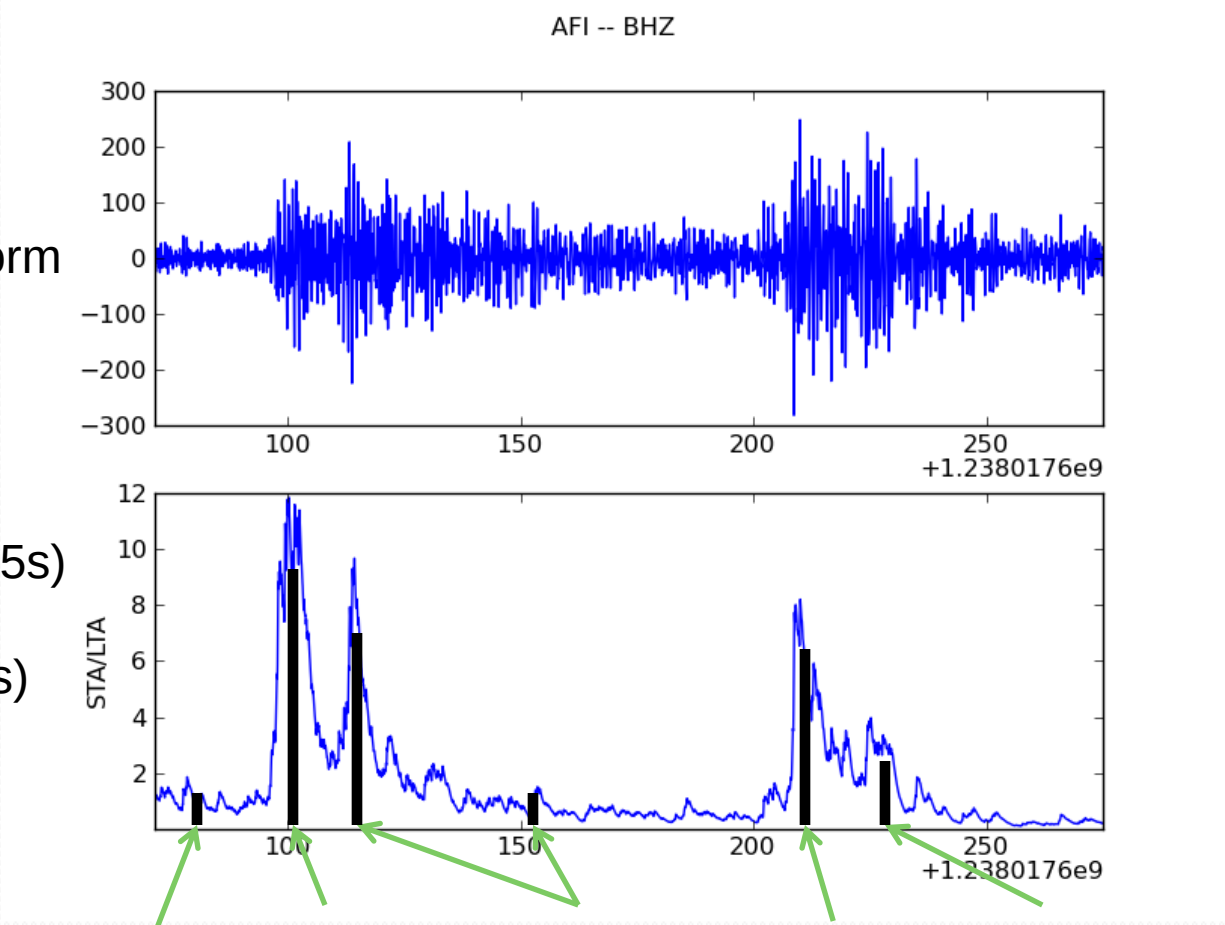


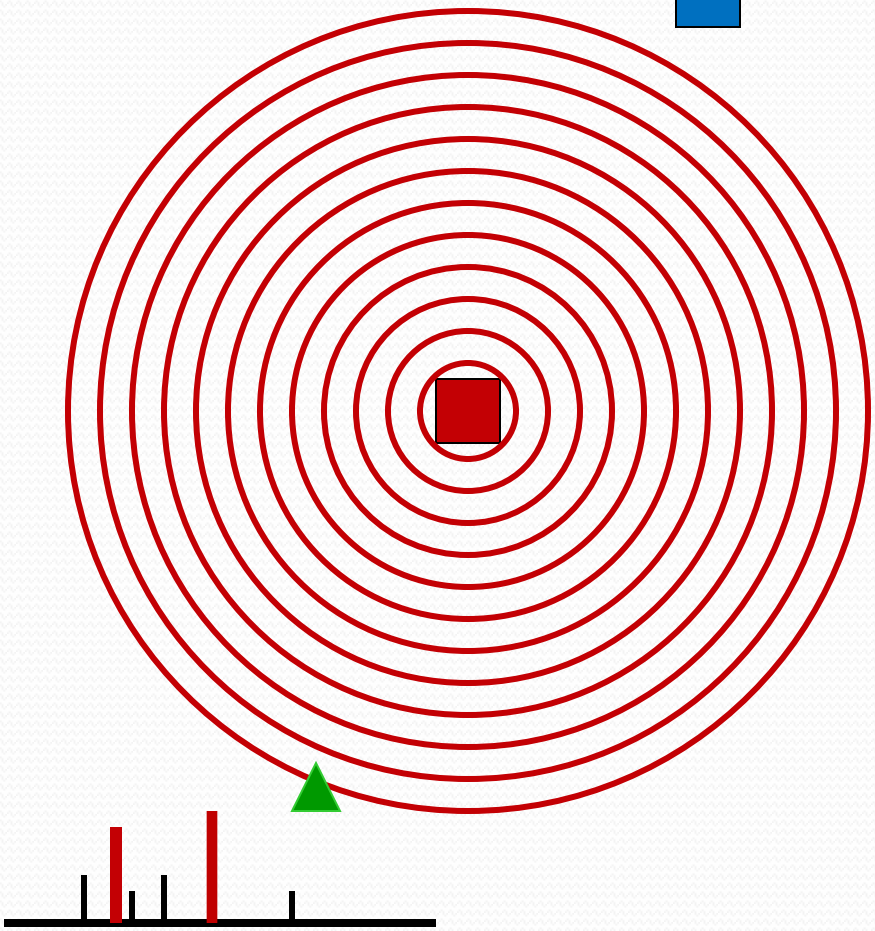


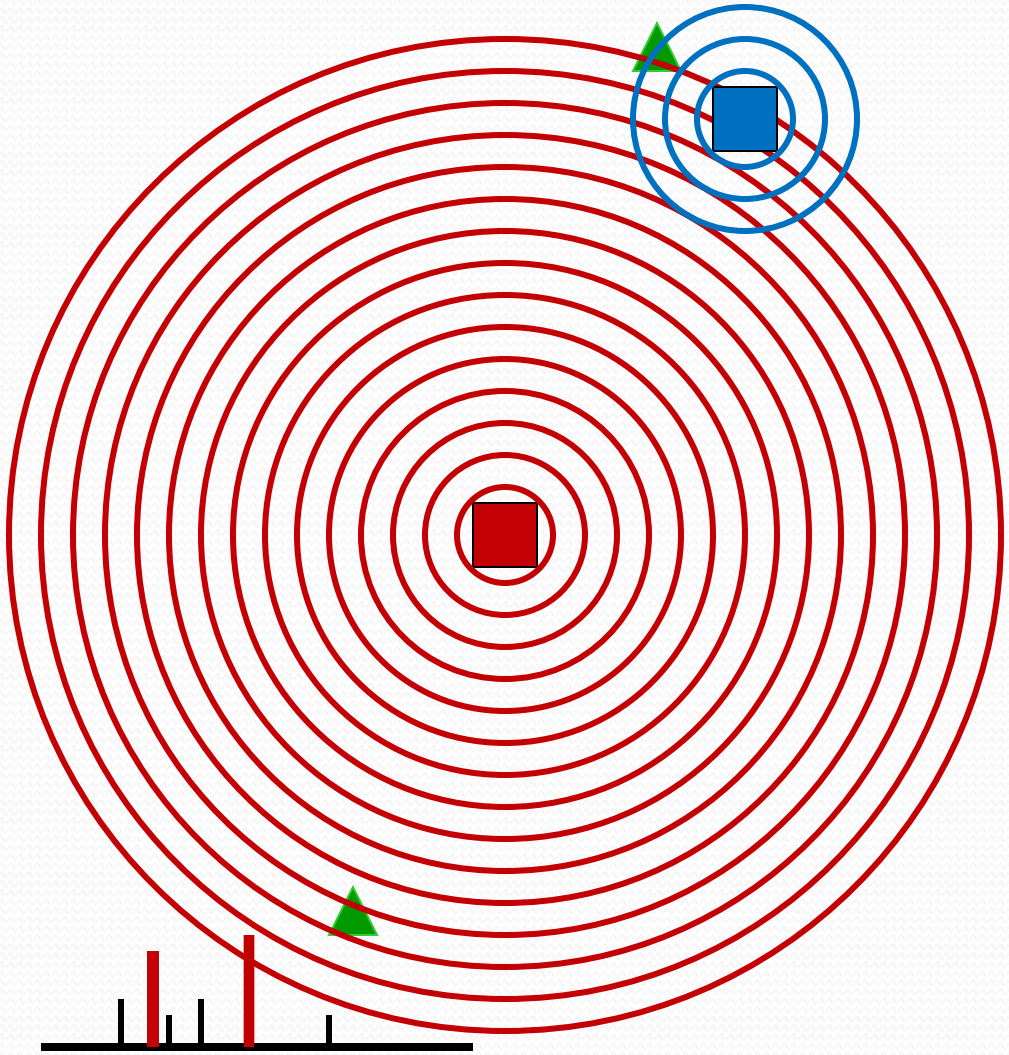
Waveforms → Detections

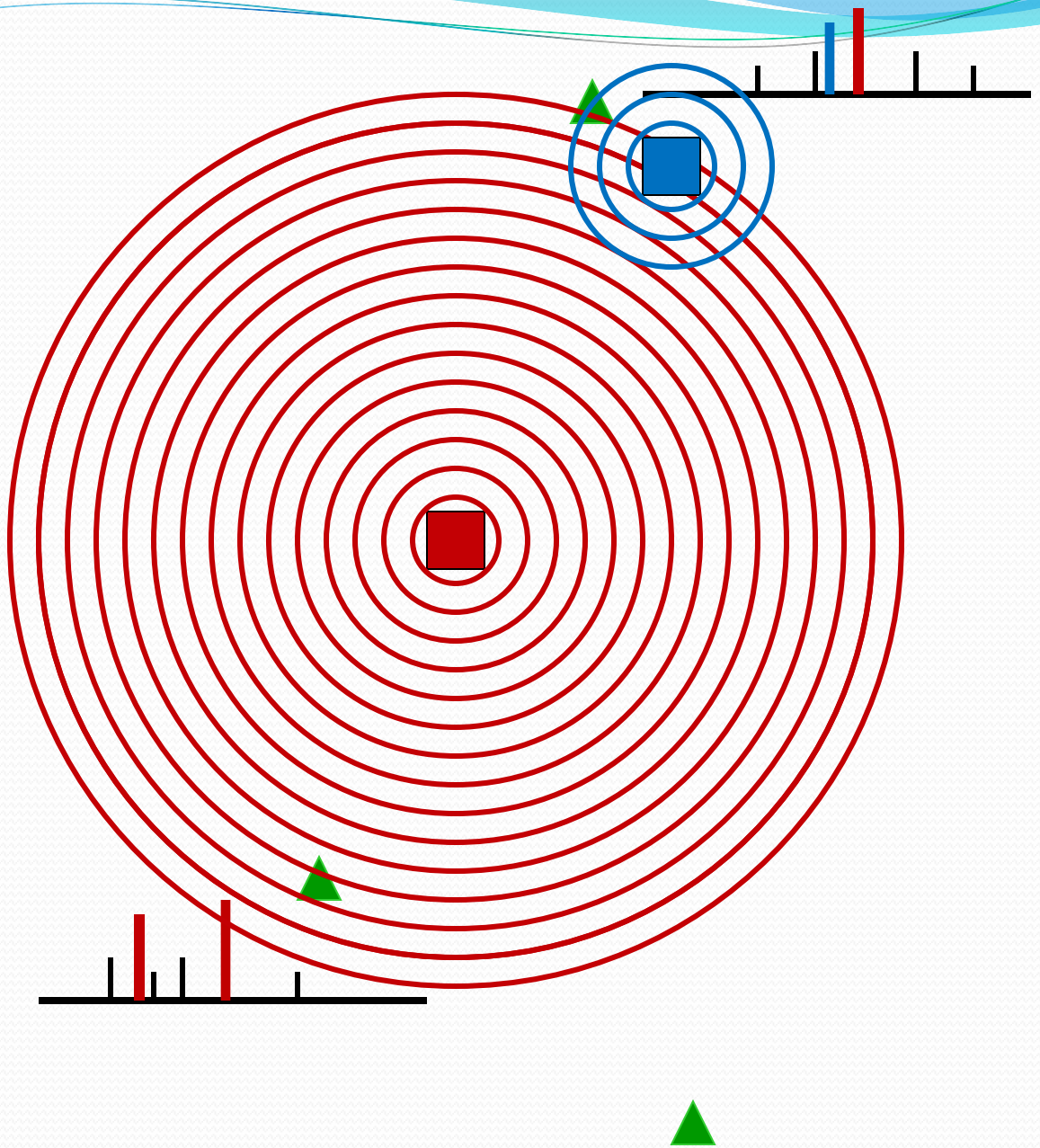
Filtered Waveform (1-4 Hz)

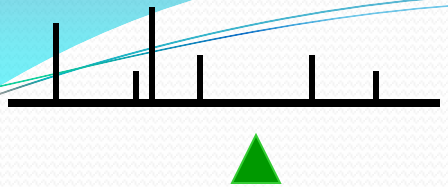
Short-term (1.5s) divided by long-term (60s) average

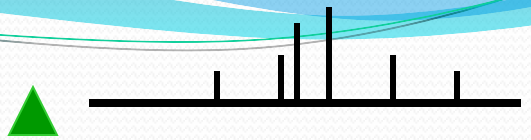


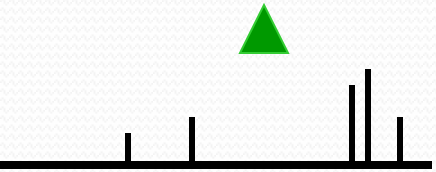


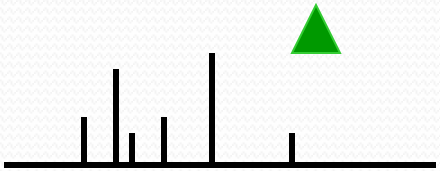


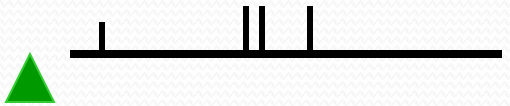












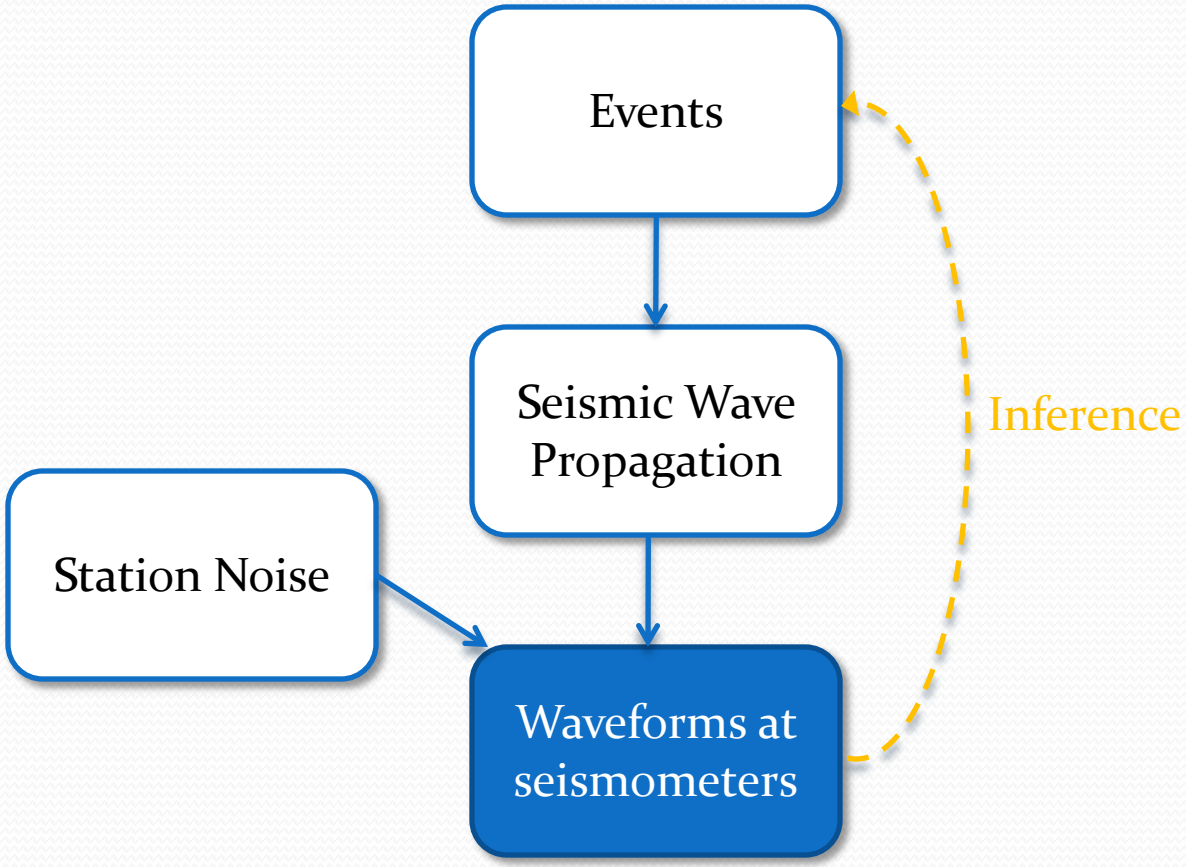
Why is the problem hard?

- ~10000 "detections" per day, 90% from noise (less than magnitude 2 events)
- CTBT needs to find ***all*** seismic events
- CTBT automated system (SEL3) finds 69% of significant events and half the predicted events are spurious (nonexistent)
- 16 human analysts are required to fix these errors, and generate LEB ("ground truth")
- Unreliable below magnitude 4 (1kT)

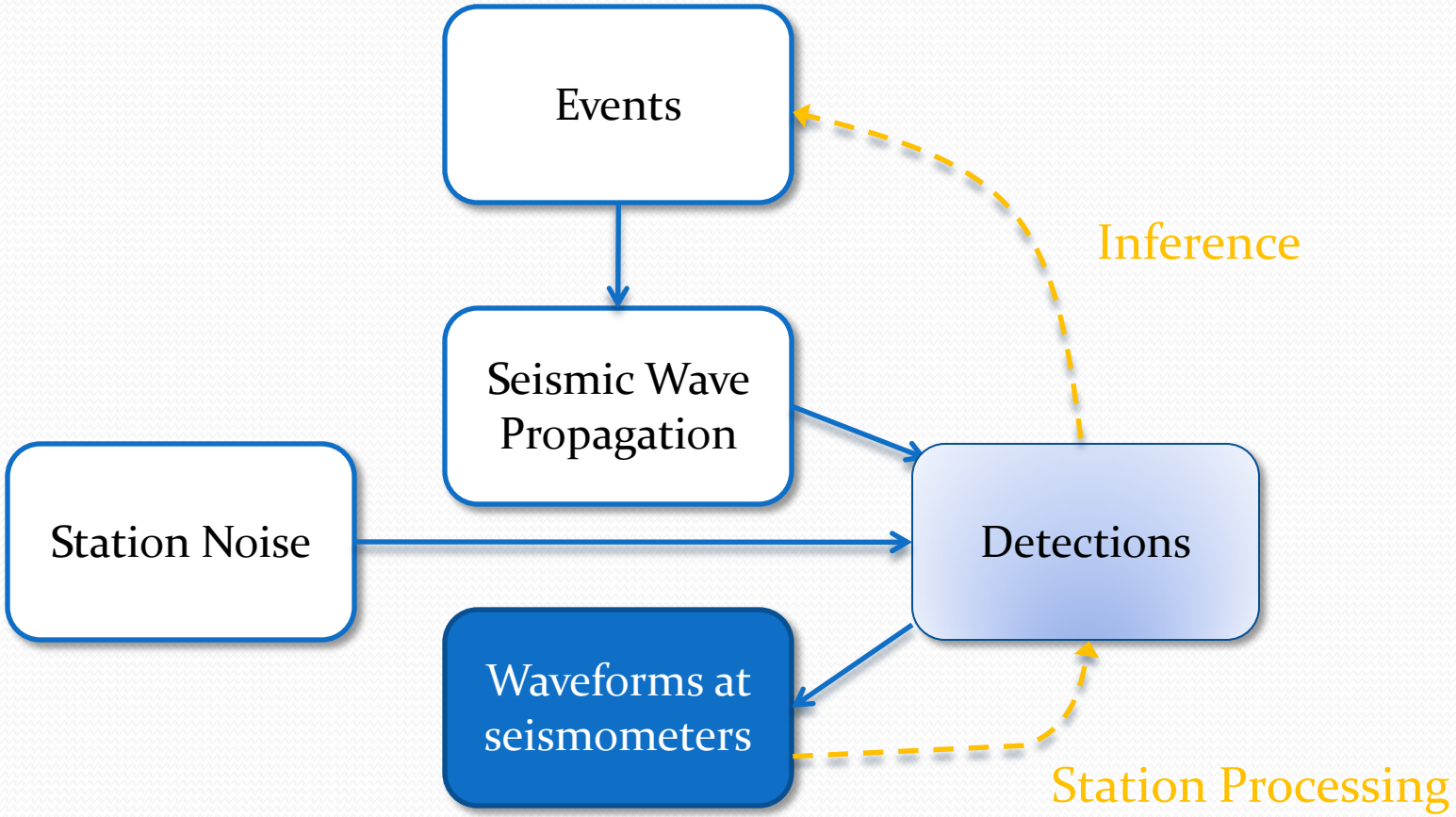
Overview

- Seismology
- Generative Model
- Inference
- Results

Vertically Integrated Seismic Analysis (VISA)



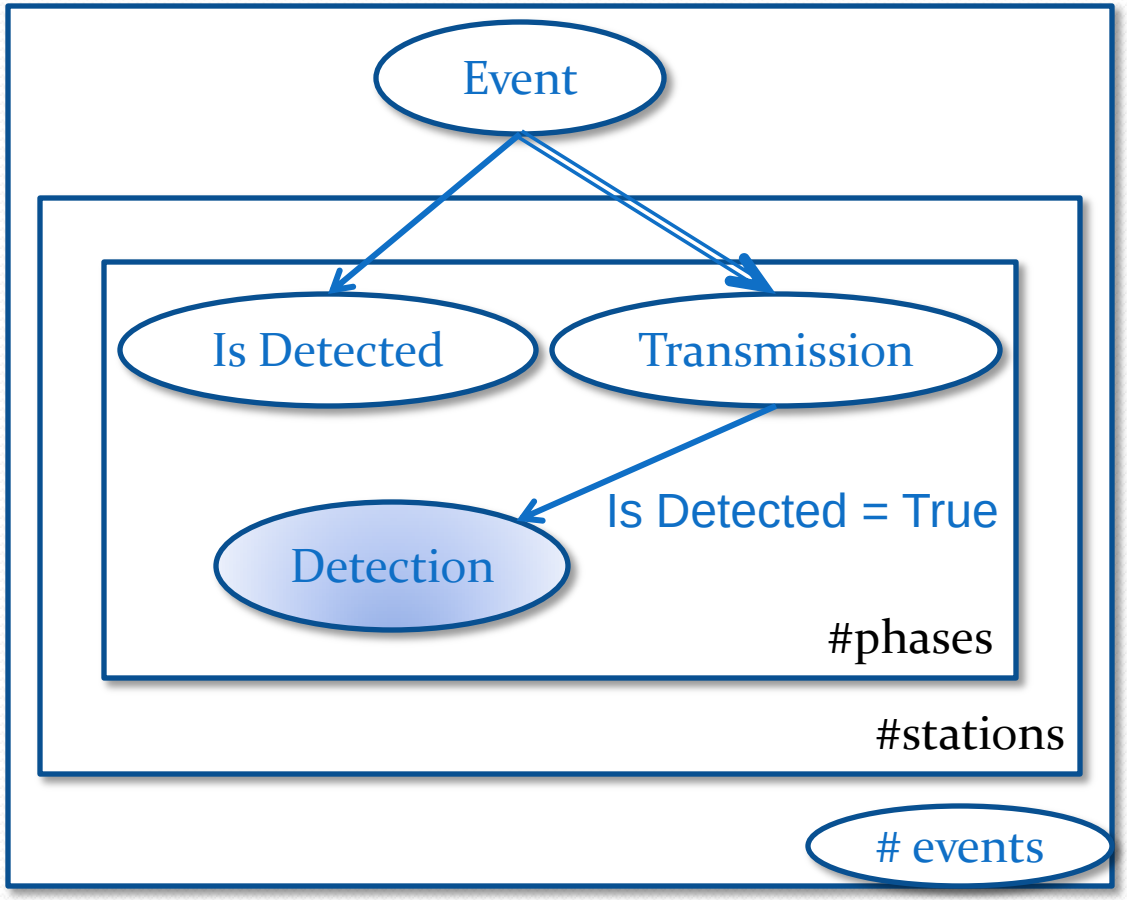
Network Processing (NET-VISA)

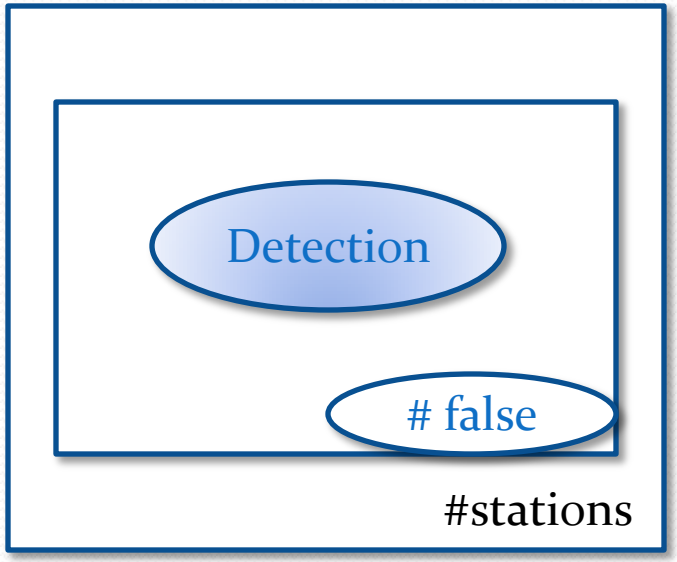


Expressed in BLOG (Bayesian Logic)

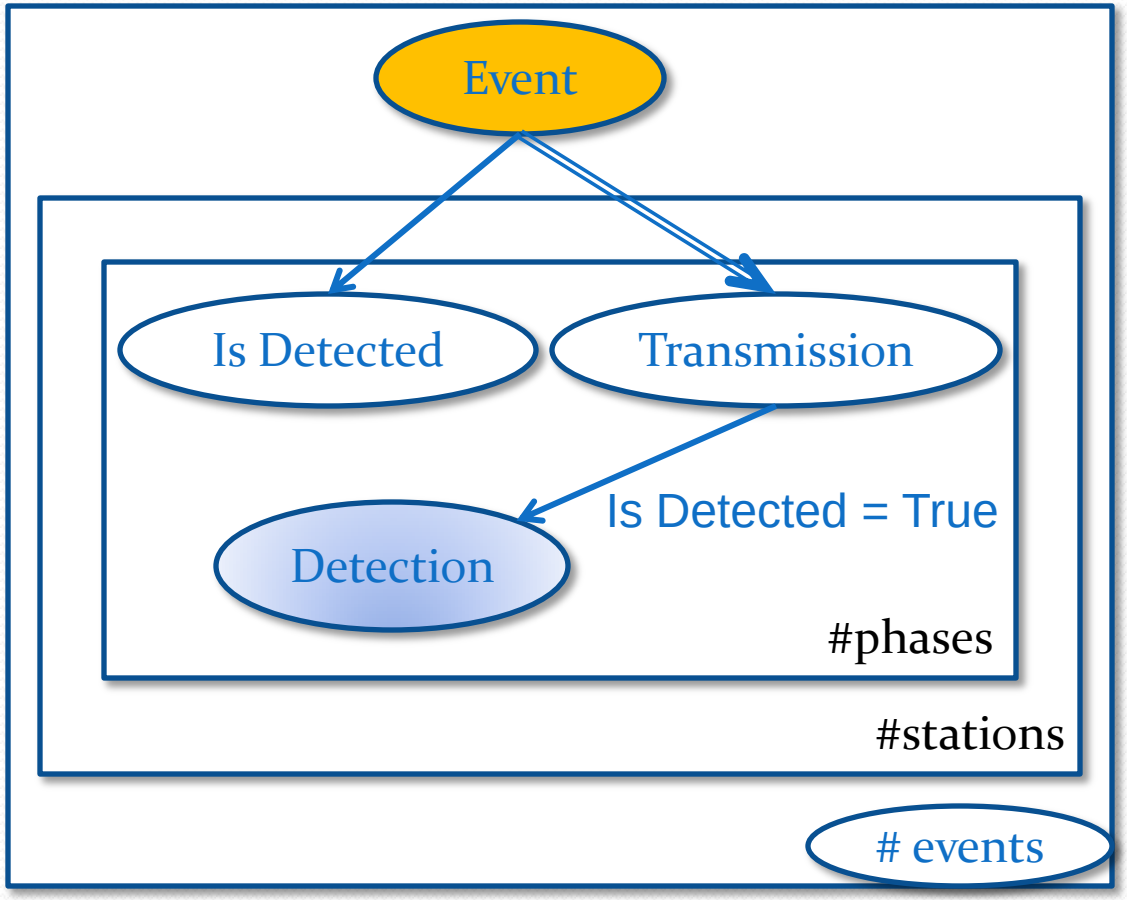
 # SeismicEvents ~ Poisson[TIME_DURATION*EVENT_RATE];
IsEarthQuake(e) ~ Bernoulli(.999);
EventLocation(e) ~ If IsEarthQuake(e) then EarthQuakeDistribution()
Else UniformEarthDistribution();
Magnitude(e) ~ Exponential(log(10)) + MIN_MAG;
Distance(e,s) = GeographicalDistance(EventLocation(e), SiteLocation(s));
IsDetected(e,p,s) ~ Logistic[SITE_COEFFS(s,p)](Magnitude(e), Distance(e,s));
#Arrivals(site = s) ~ Poisson[TIME_DURATION*FALSE_RATE(s)];
#Arrivals(event=e, site=s) = If IsDetected(e,s) then 1 else 0;
Time(a) ~ If (event(a) = null) then Uniform(0,TIME_DURATION)
else IASPEI(EventLocation(event(a)),SiteLocation(site(a)),Phase(a)) + TimeRes(a);
TimeRes(a) ~ Laplace(TIMLOC(site(a)), TIMSCALE(site(a)));
Azimuth(a) ~ If (event(a) = null) then Uniform(0, 360)
else = GeoAzimuth(EventLocation(event(a)),SiteLocation(site(a)) + AzRes(a);
AzRes(a) ~ Laplace(AZLOC(site(a)) AZSCALE(site(a)));
Slow(a) ~ If (event(a) = null) then Uniform(0,20)
else = IASPEI-SLOW(EventLocation(event(a)),SiteLocation(site(a)) + SlowRes(site(a));
SloRes(a) ~ Laplace(SLOLOC(site(a)), SLOSCALE(site(a)));

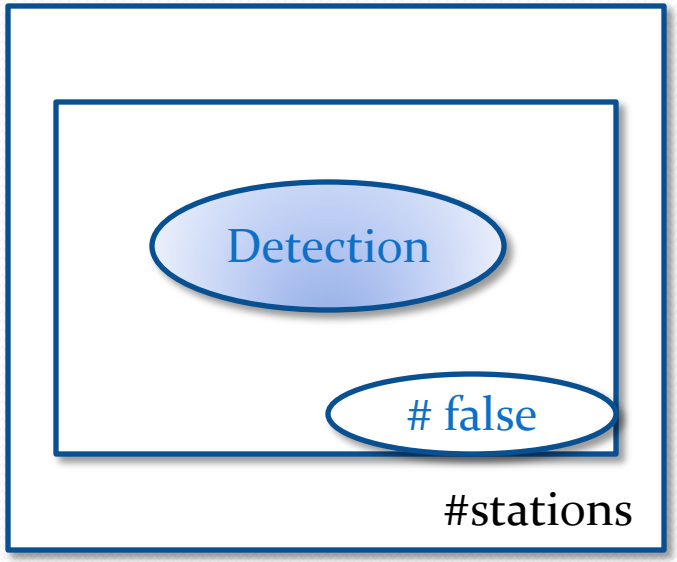
Generative Model



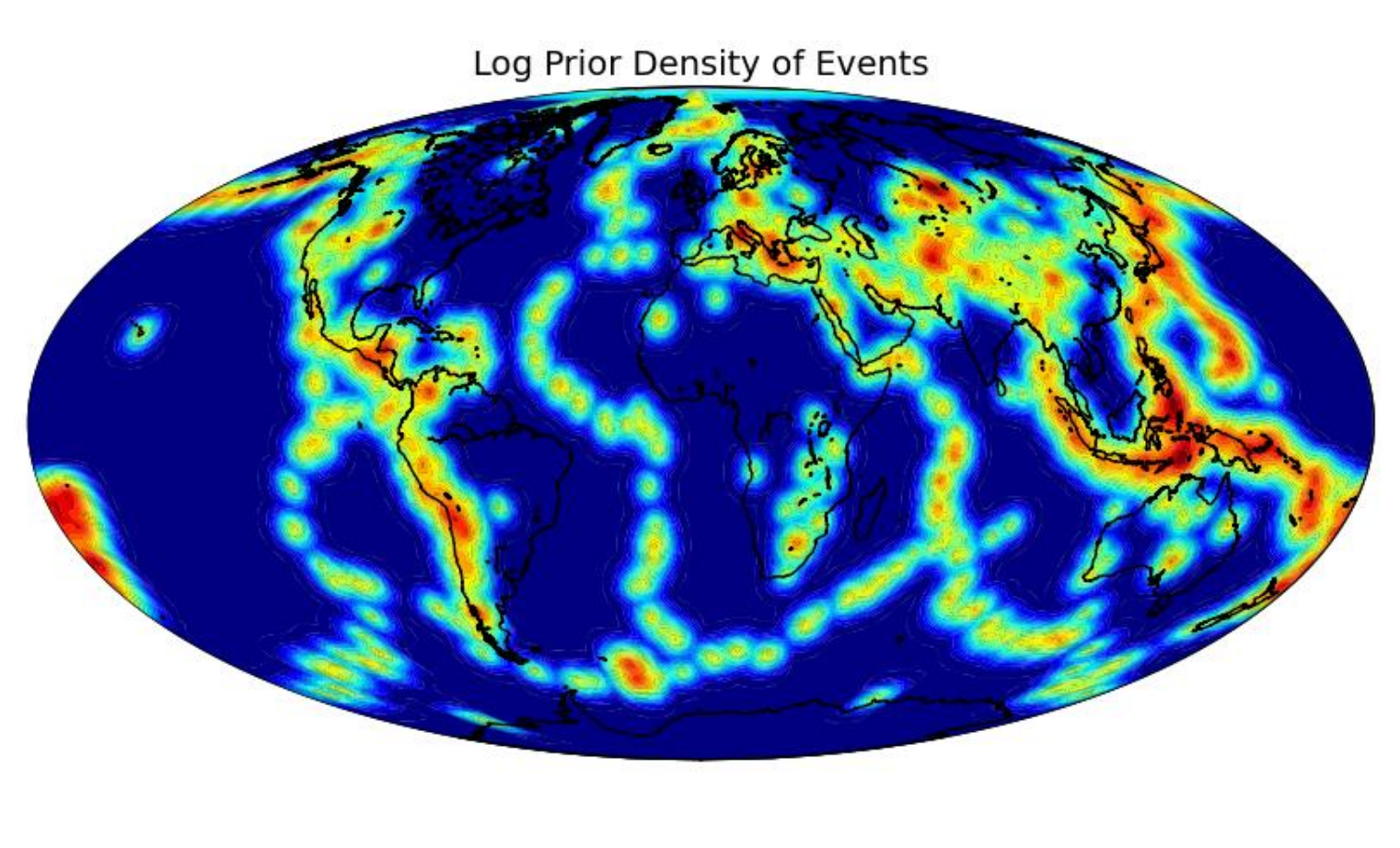


Generative Model

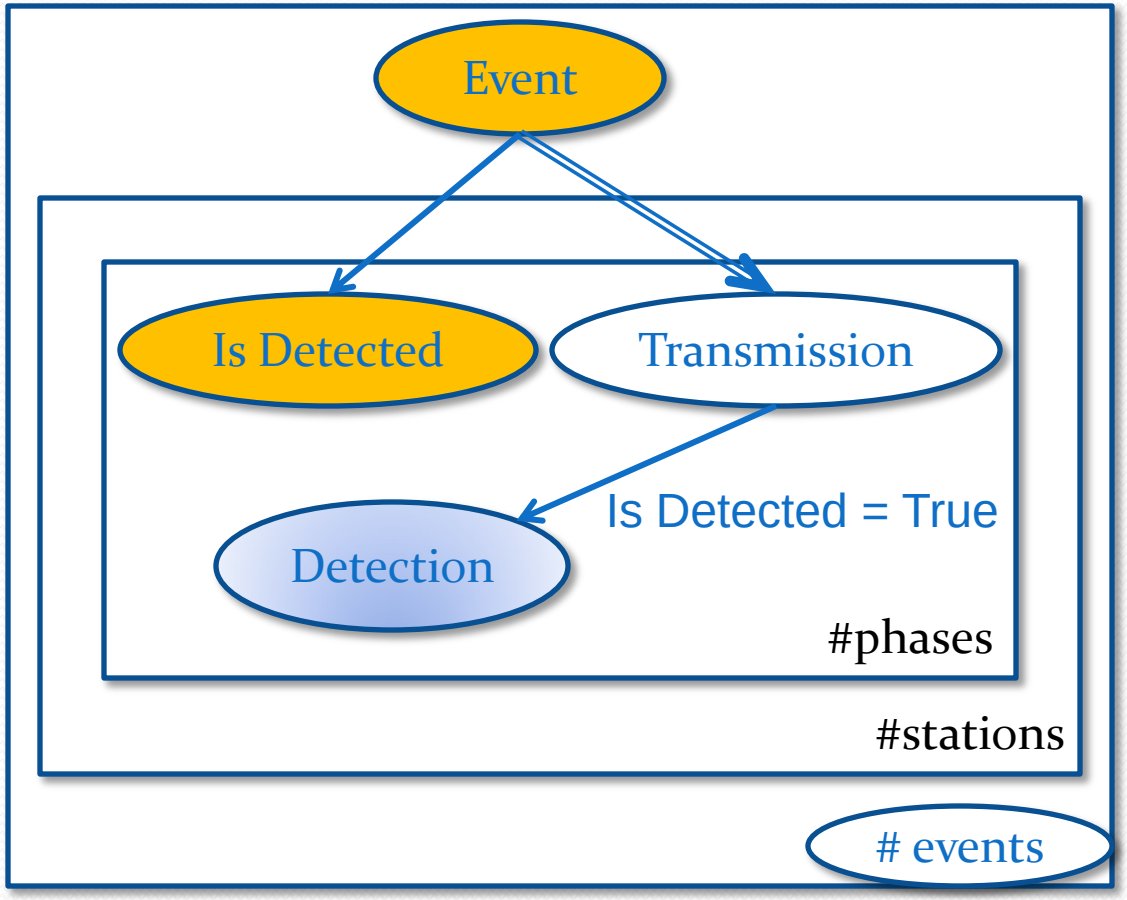


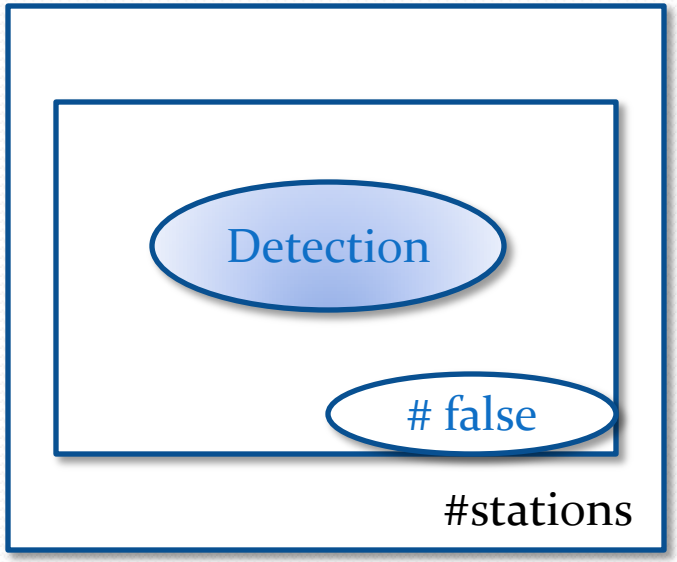


Generative Model – Event Location

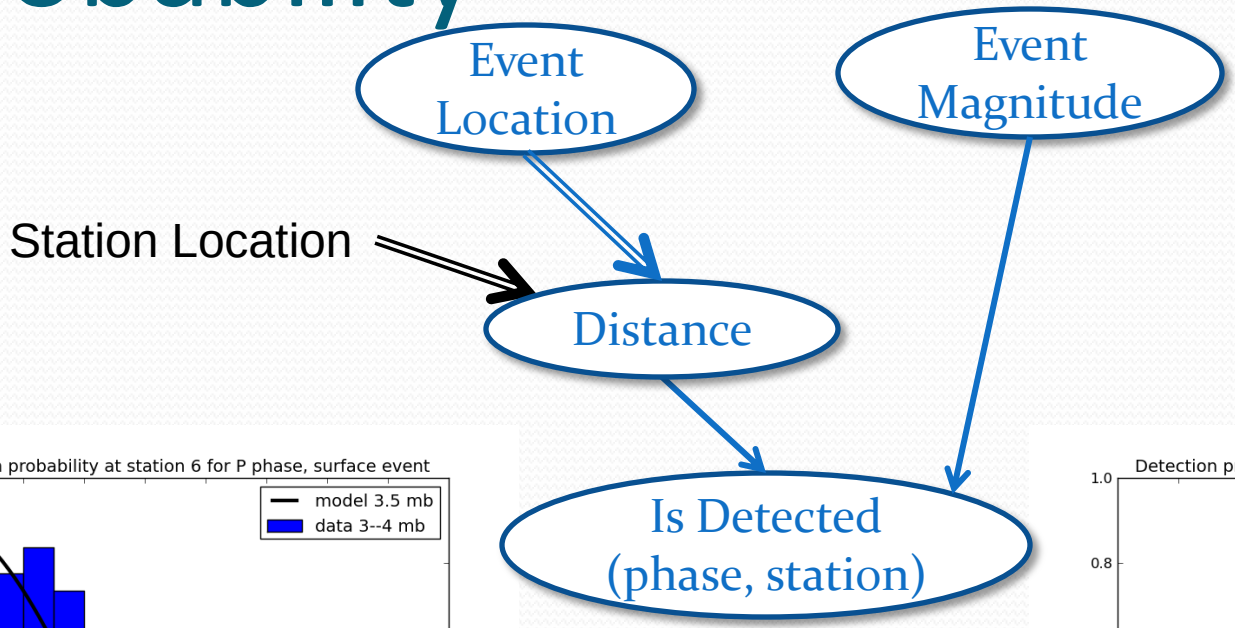


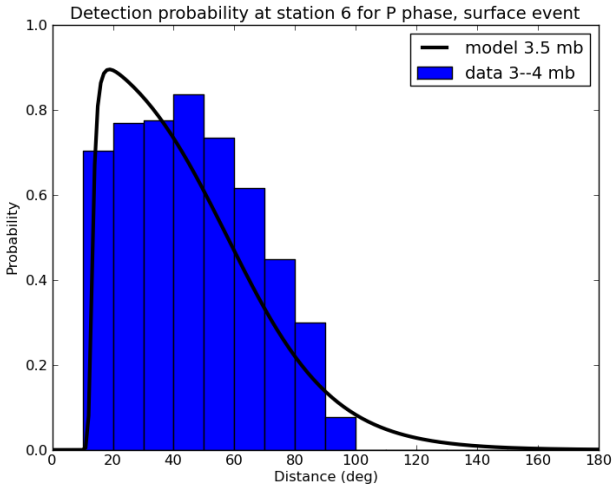
Generative Model



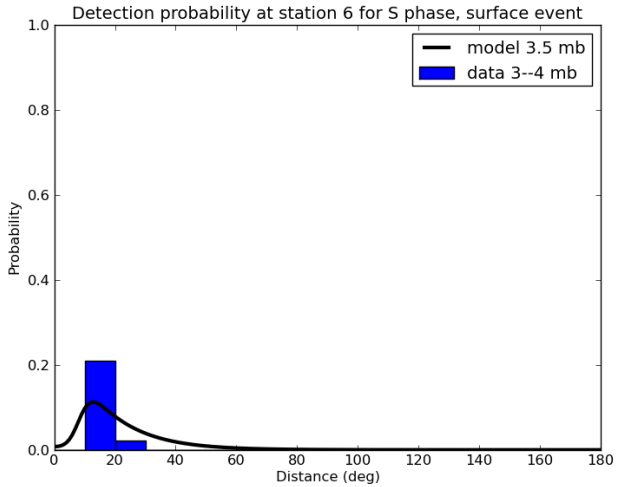


Generative Model – Detection Probability



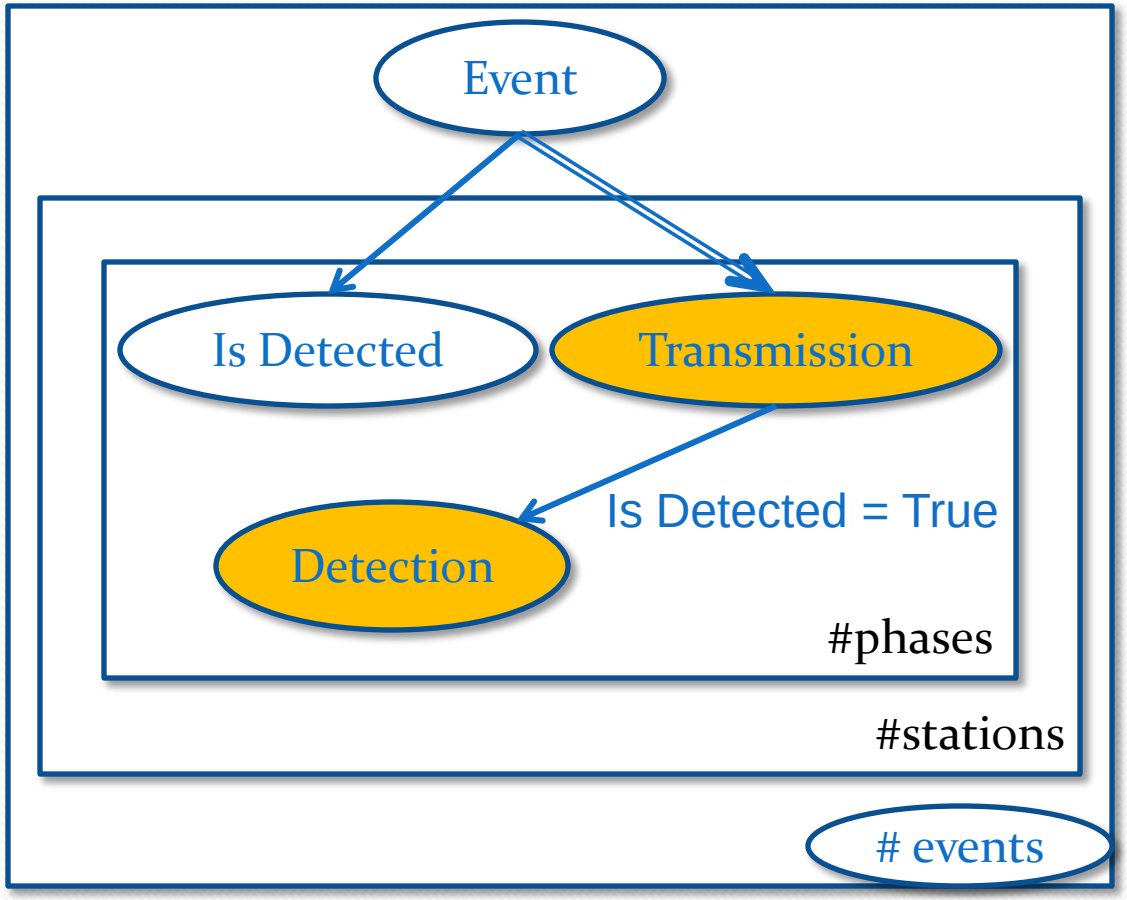


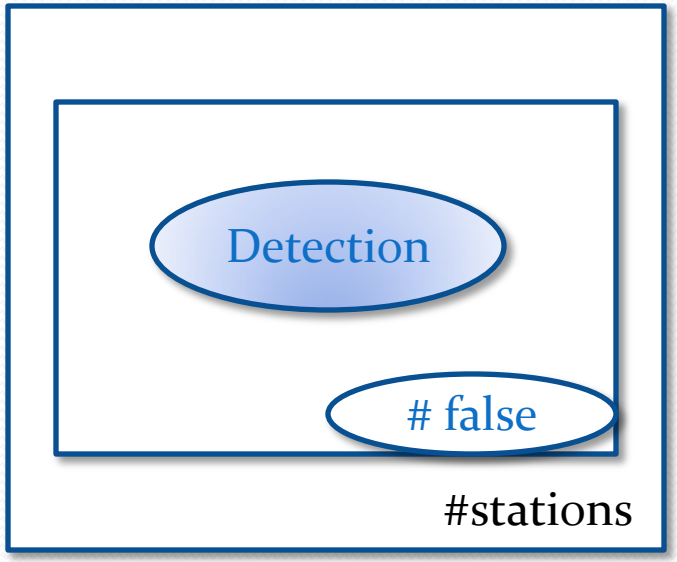
P phase, station 6



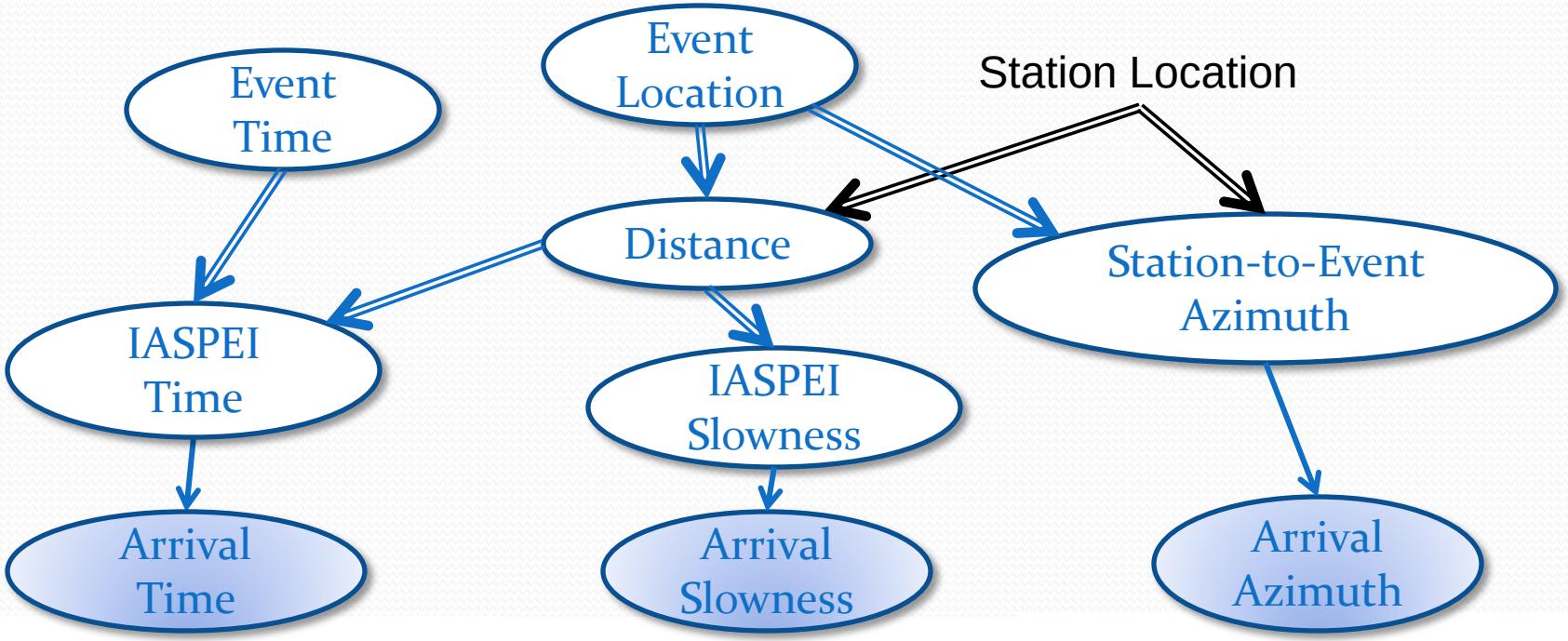
S phase, station 6

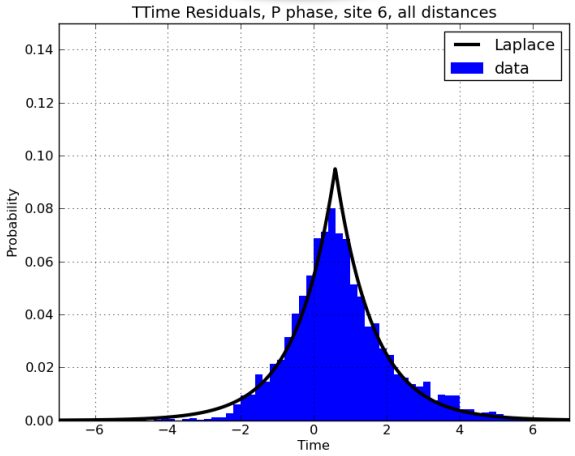
Generative Model

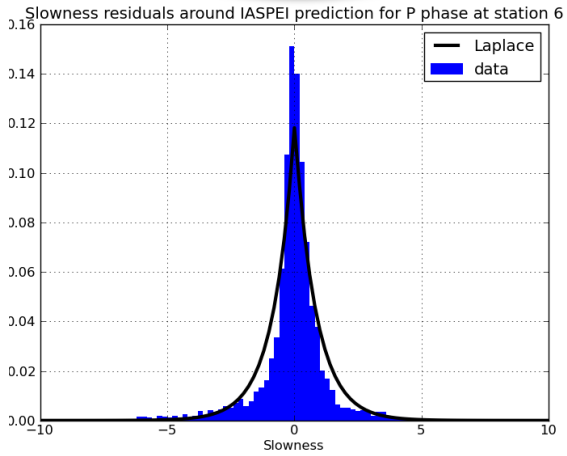


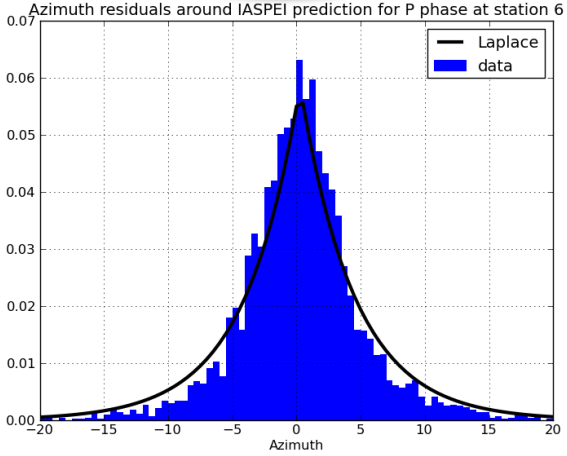


Generative Model – Arrival Time, Azimuth, Slowness

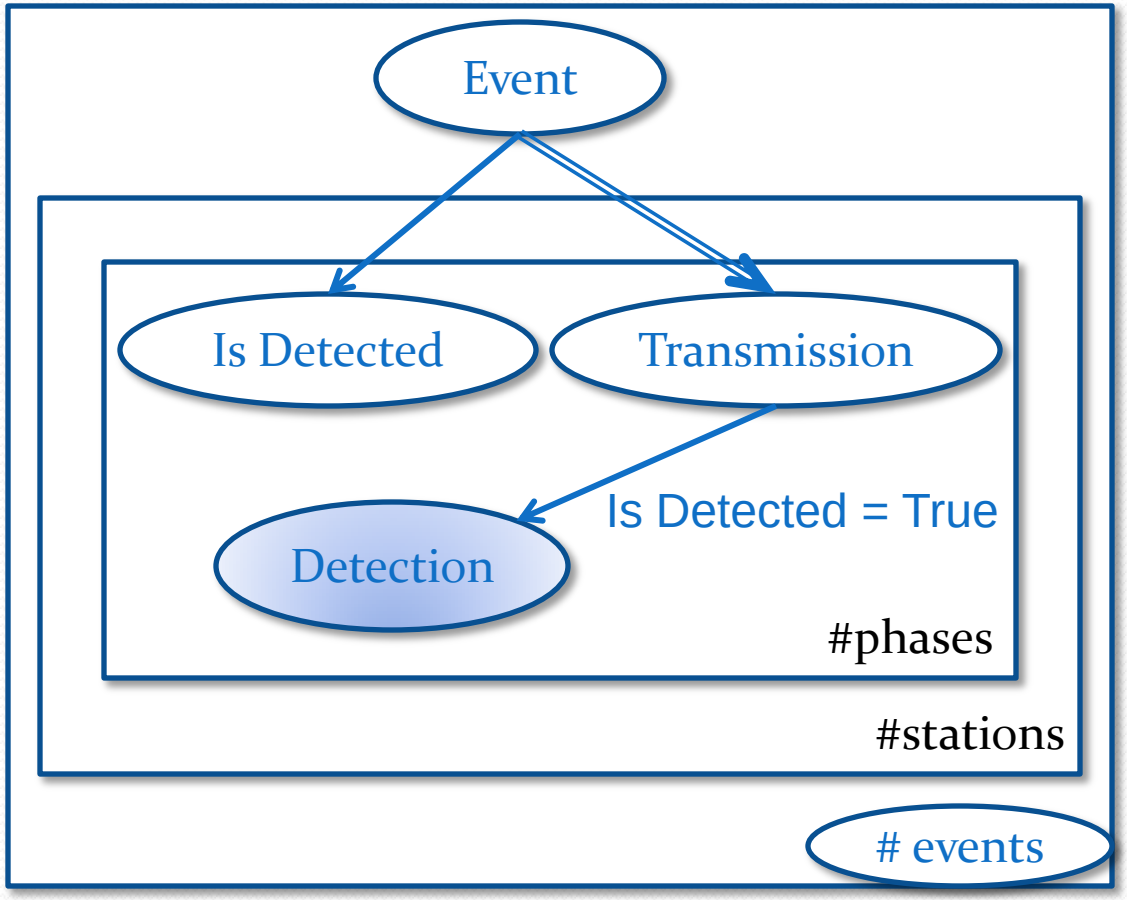


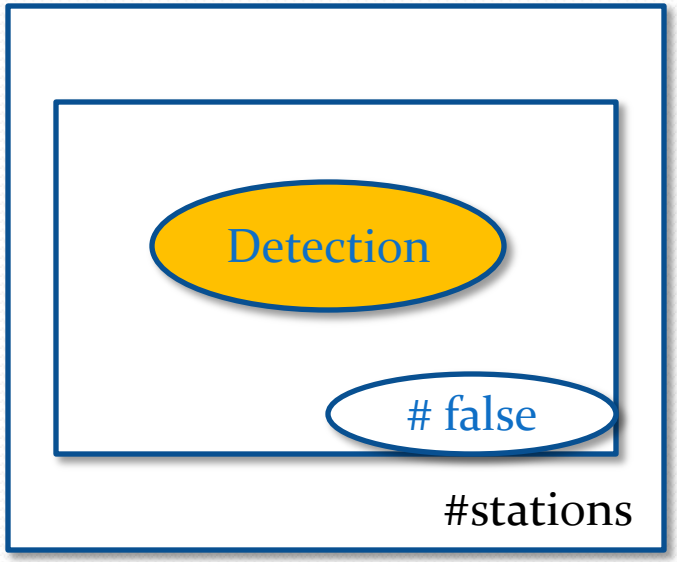




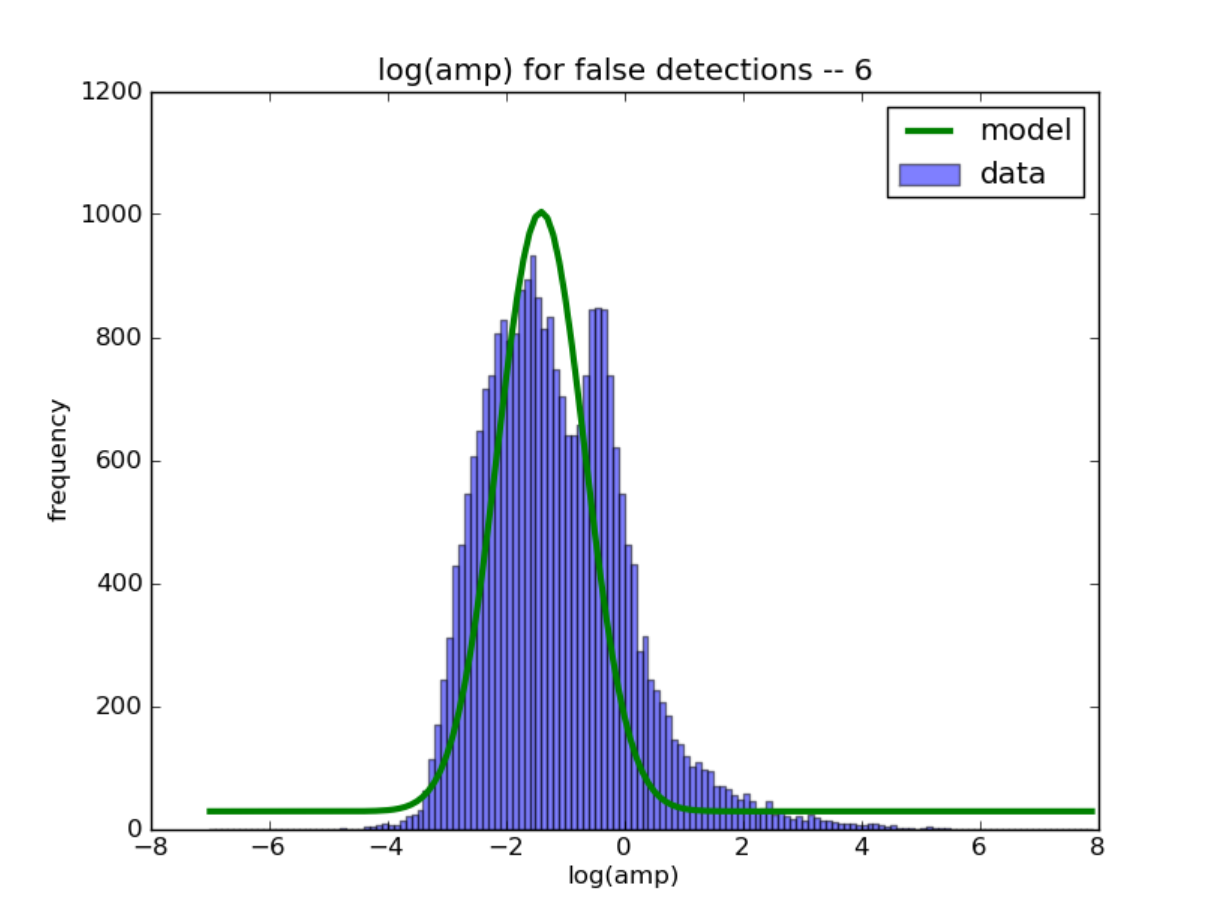


Generative Model





Noise Amplitude & Rate



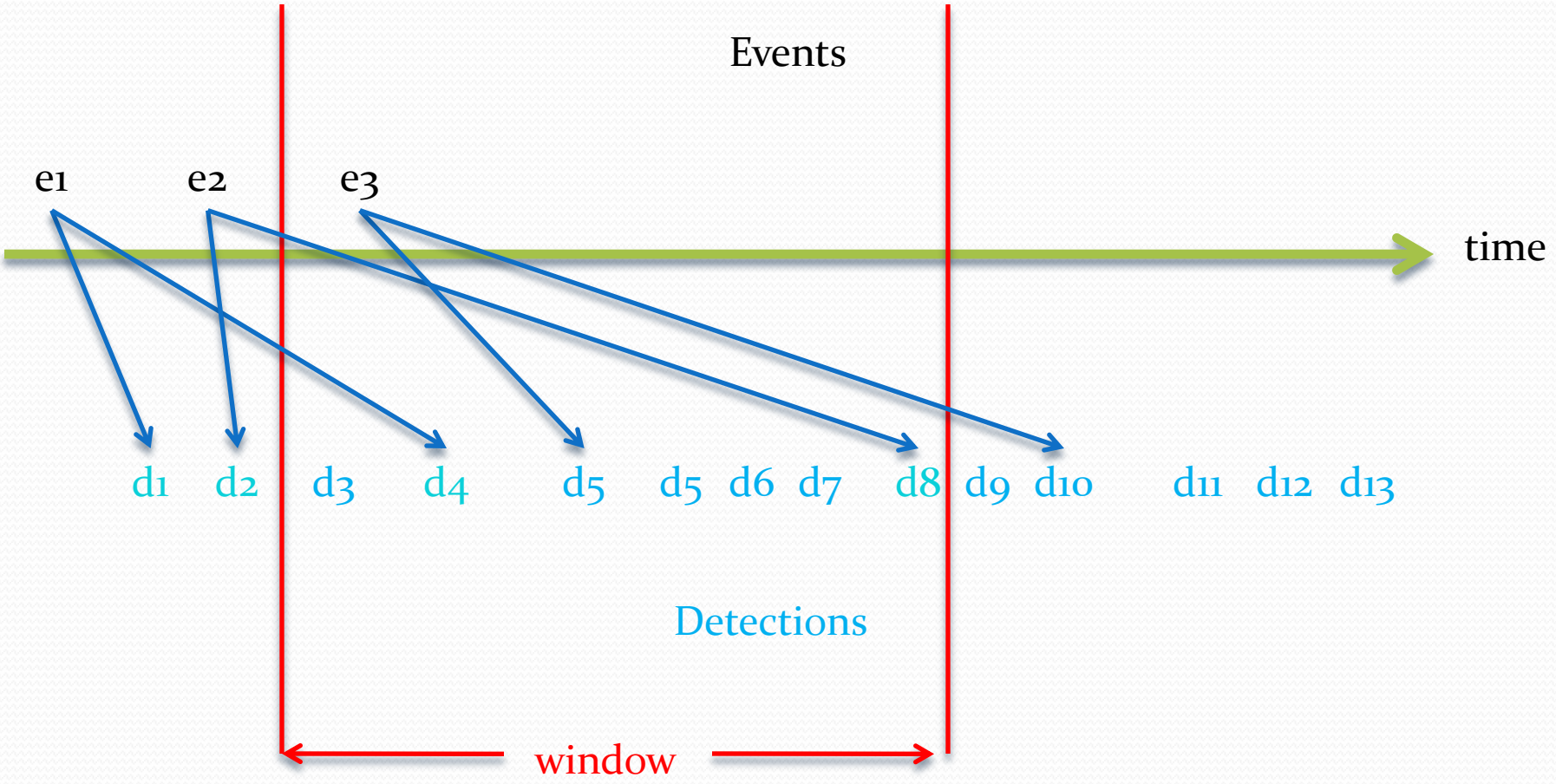
Overview

- Seismology
- Generative Model
- Inference
- Results

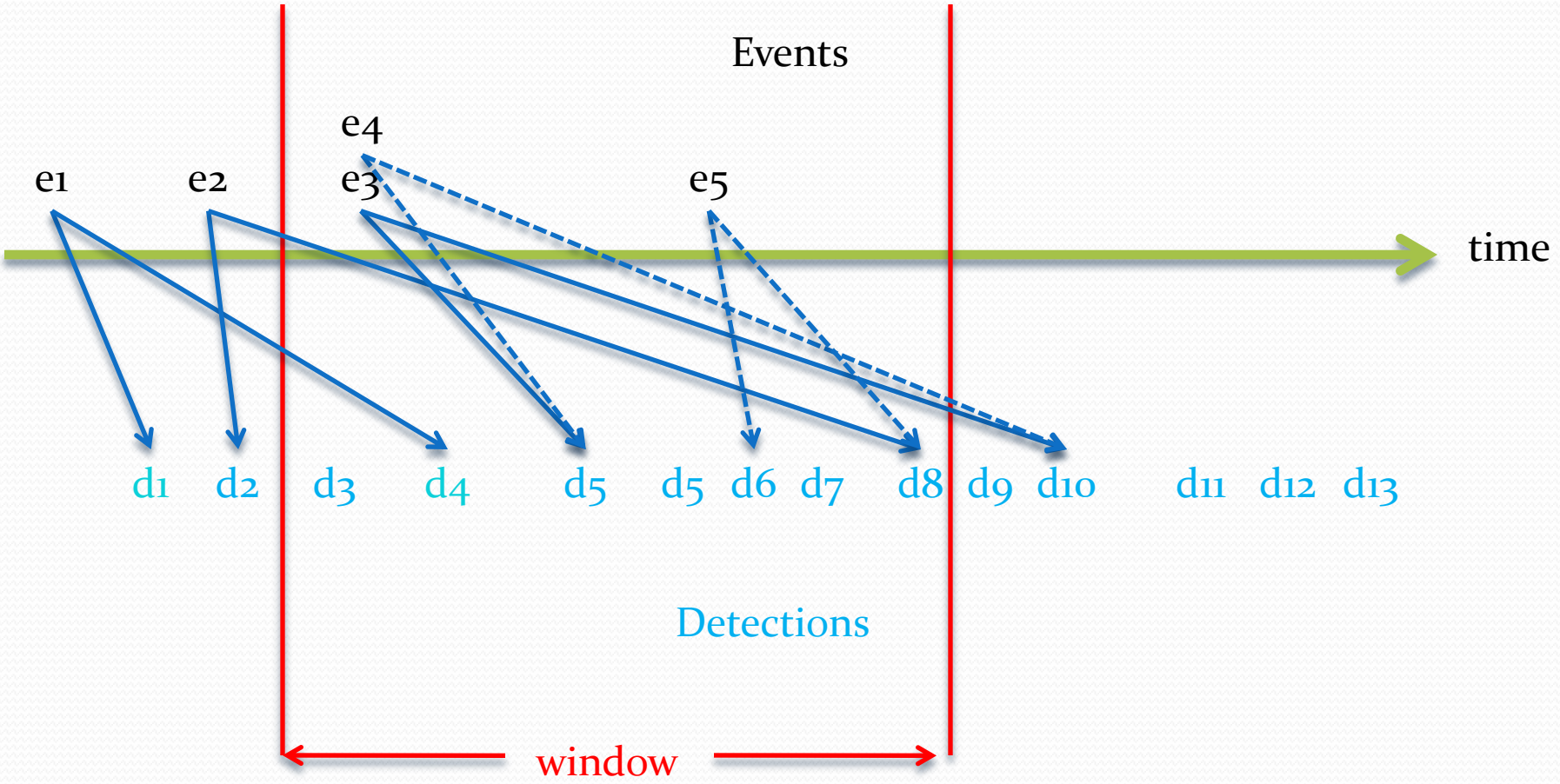
Inference Overview

- *Max a-posteriori* (MAP) seismic event bulletin
- Easier to compare to SEL3
- Continuously incorporate new detections in the hypothesis
- Heuristic search moves improve the probability
  - Birth
  - Reassociate
  - Relocate
  - Death

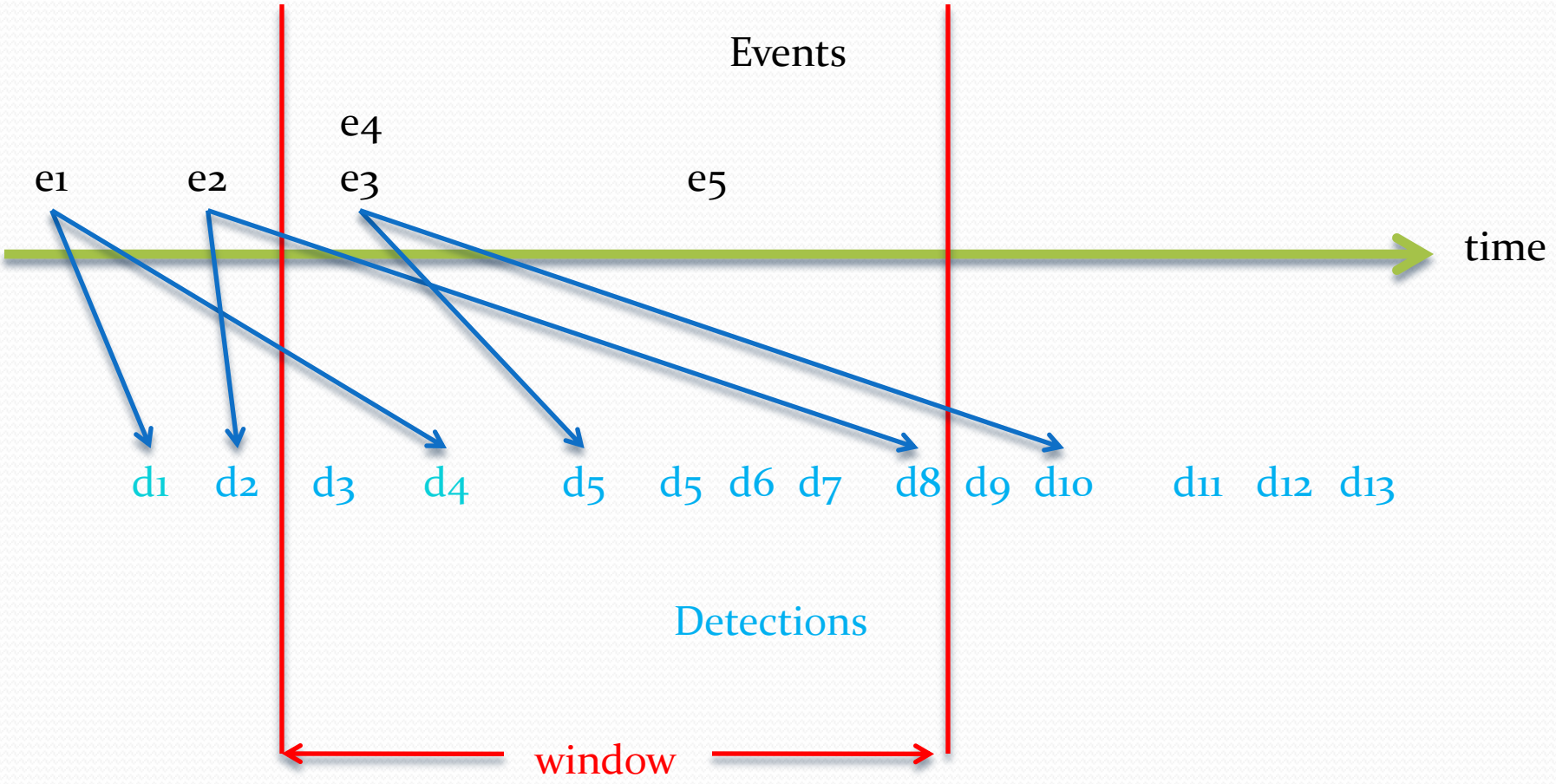
Inference Example



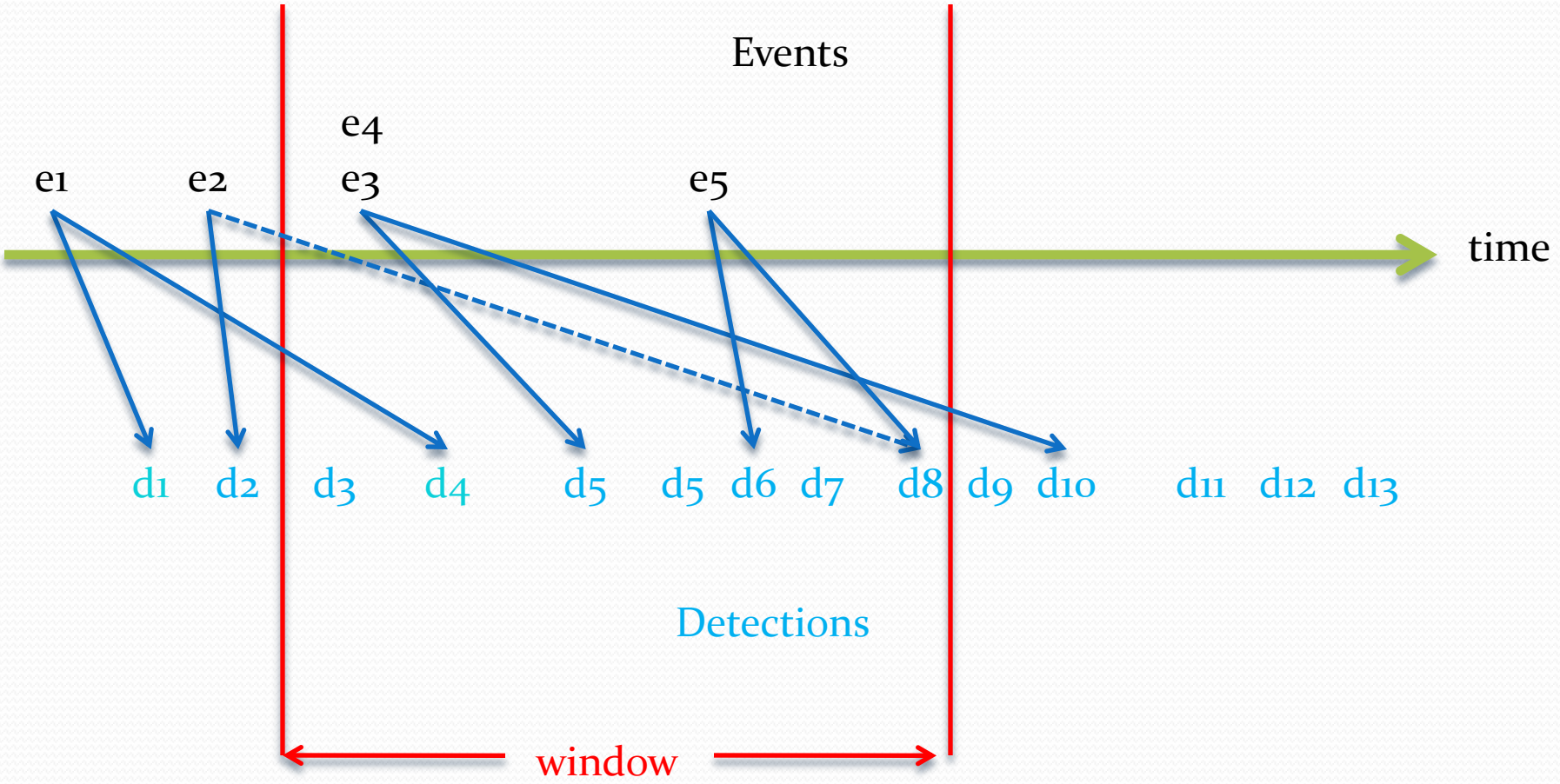
Inference : Birth Move



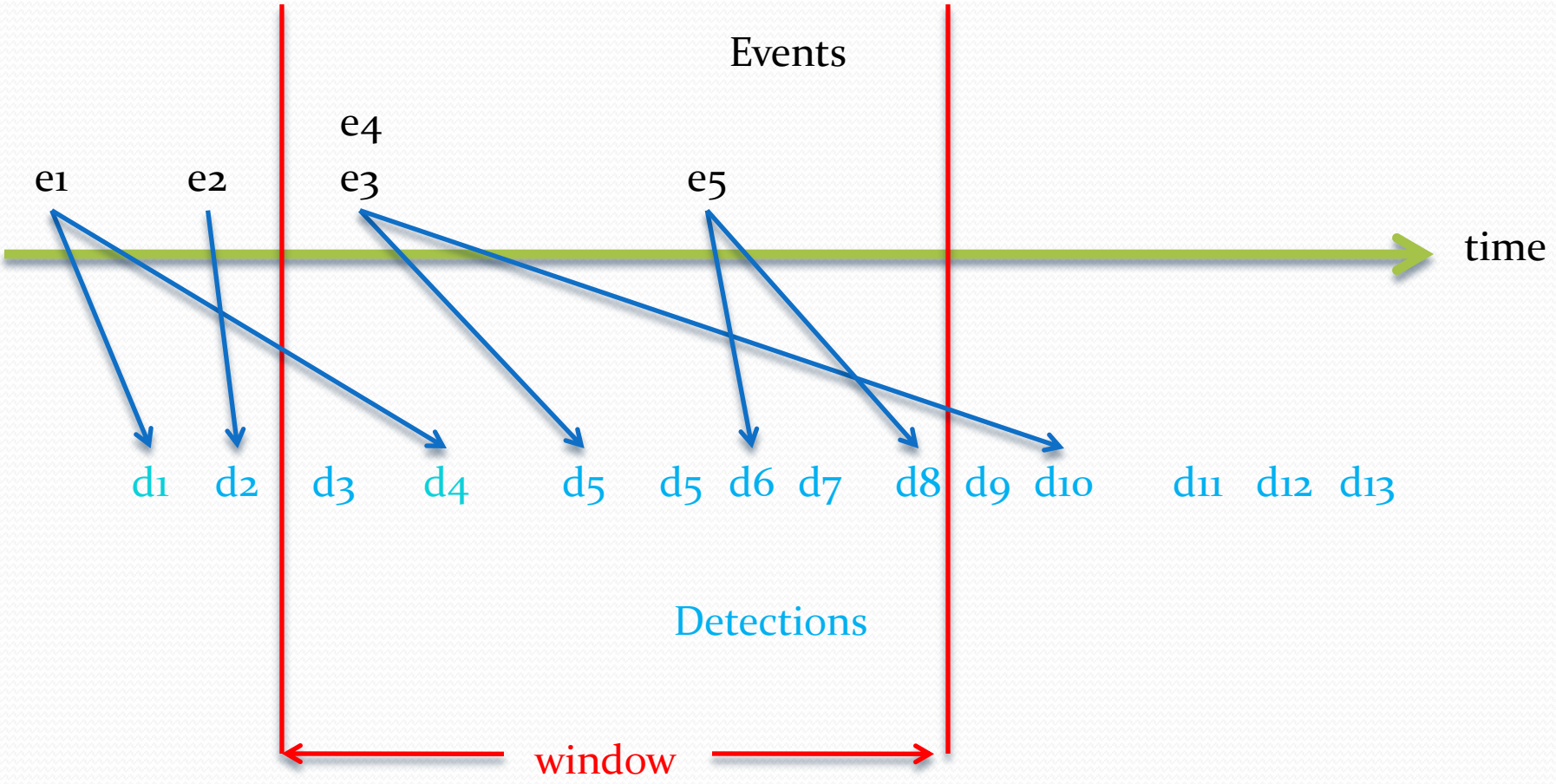
Inference : Birth Move



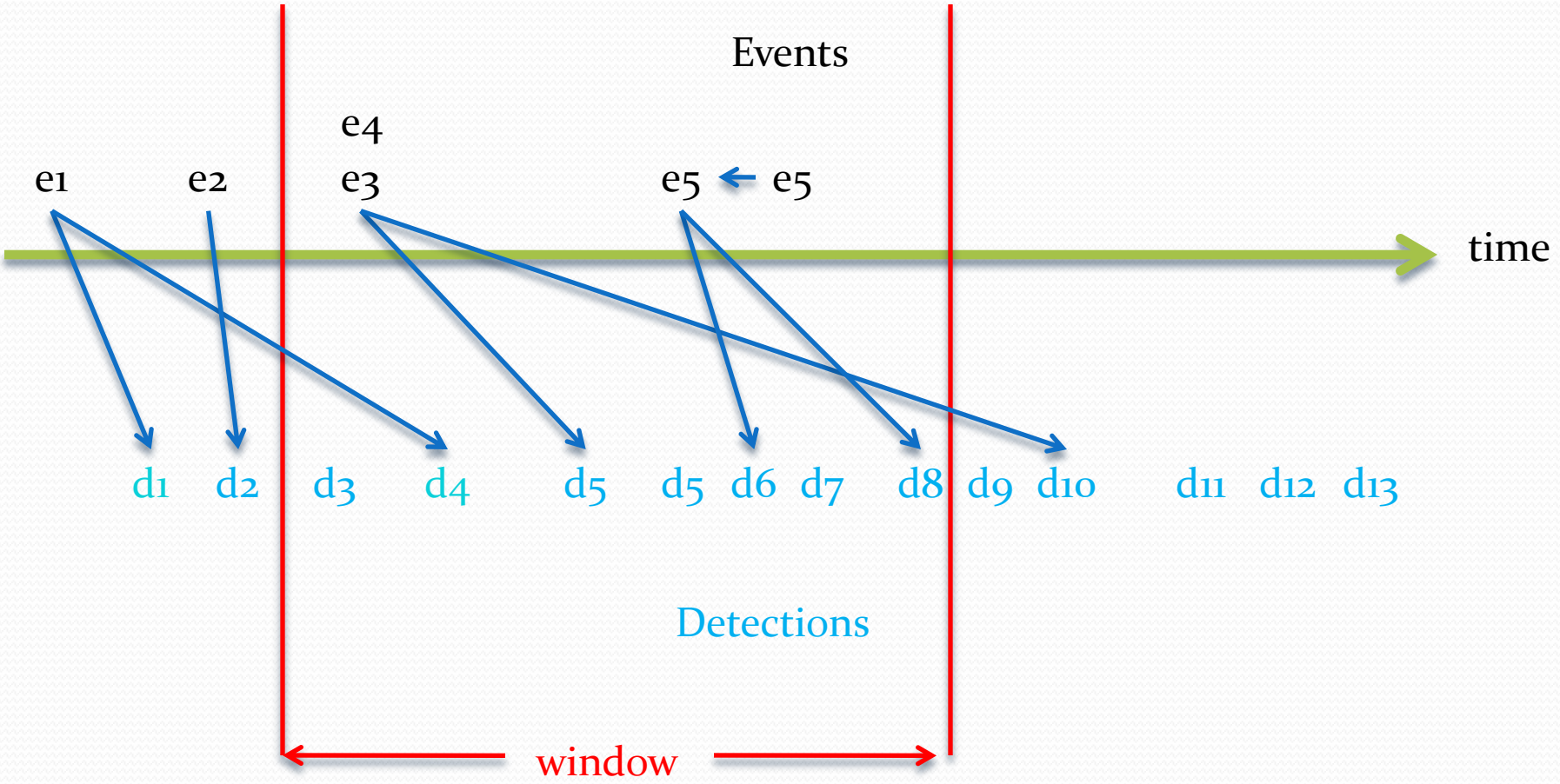
Inference : Reassociate Detections



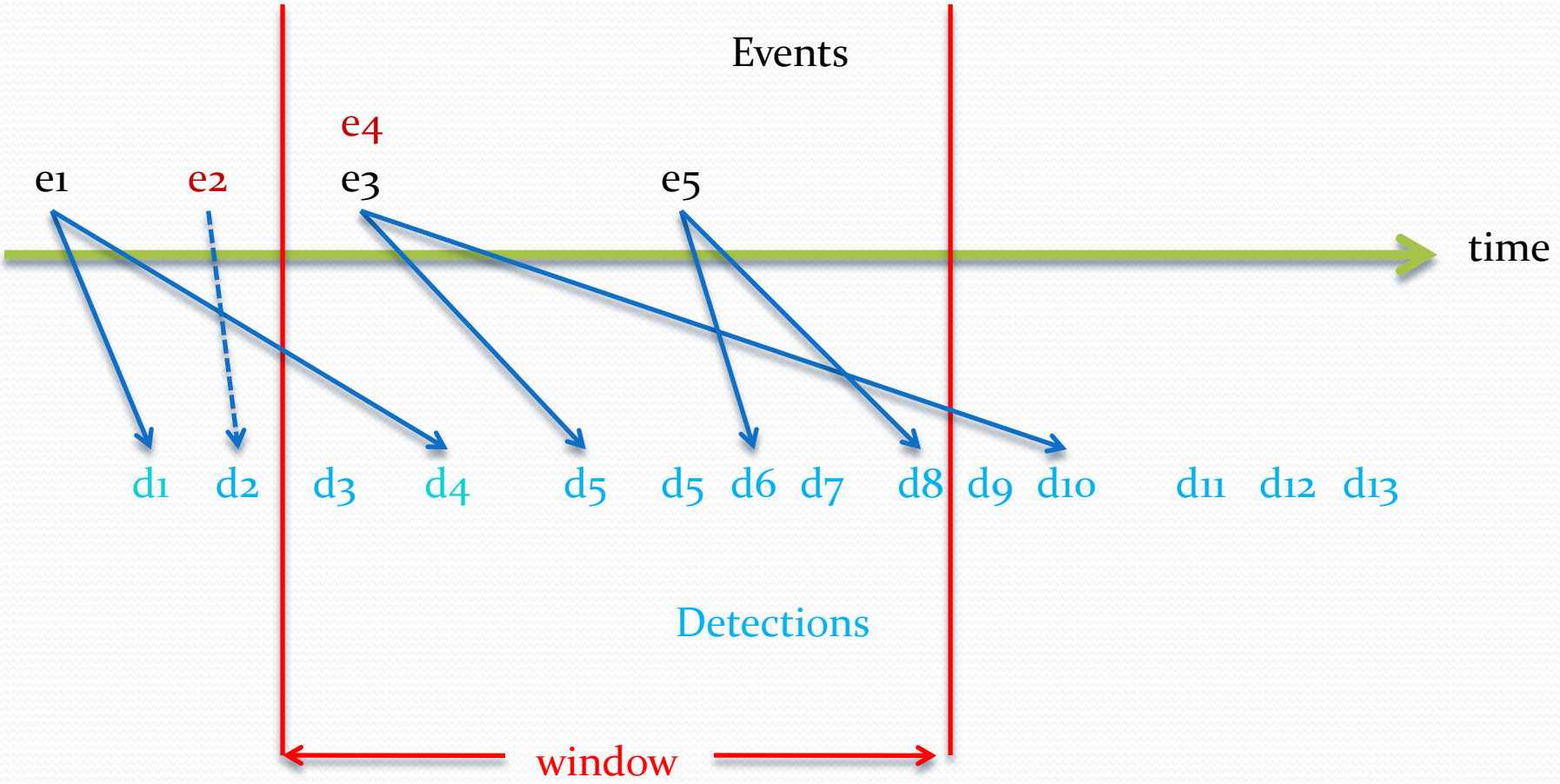
Inference : Reassociate Detections



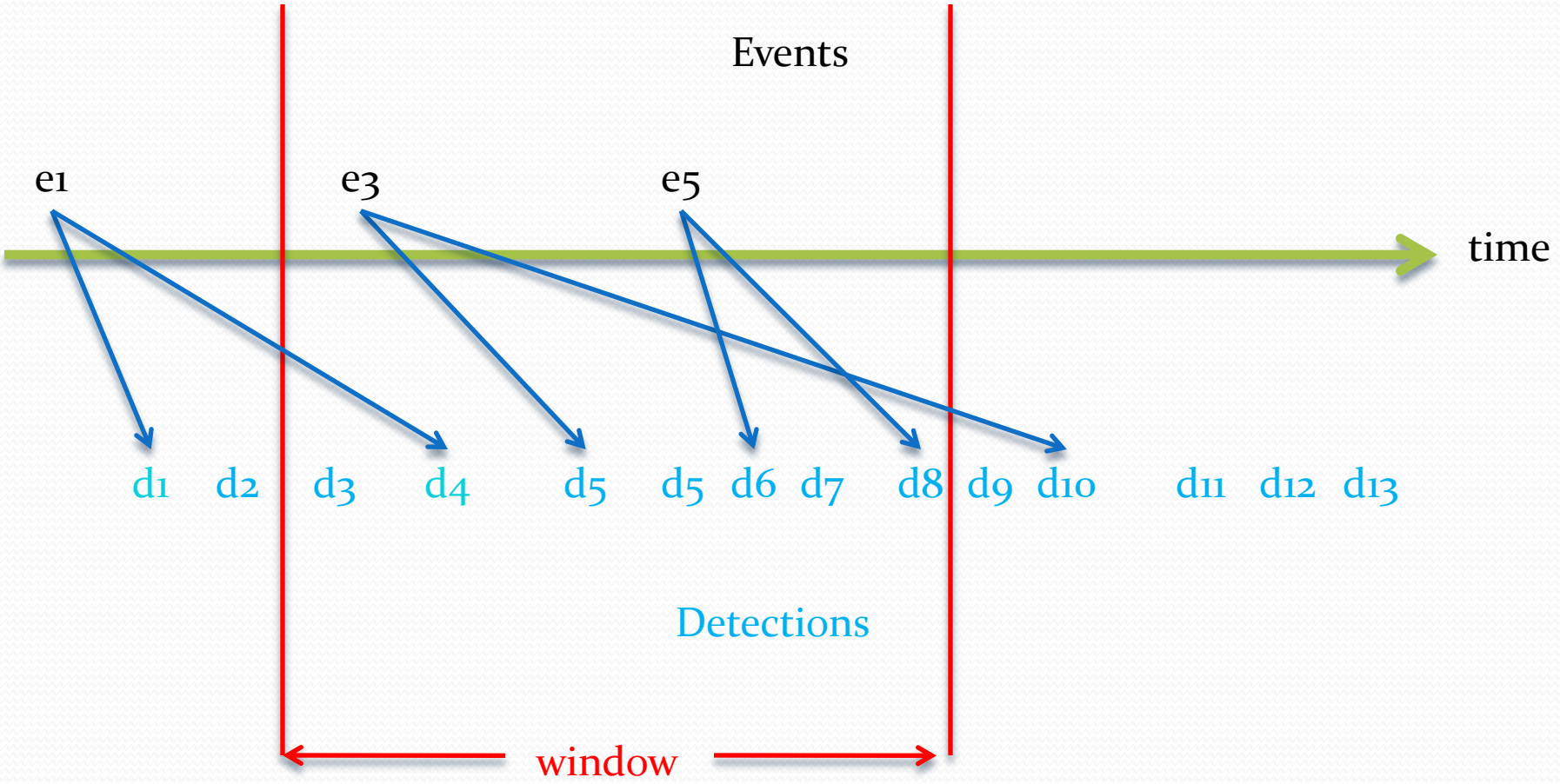
Inference : Relocate Events



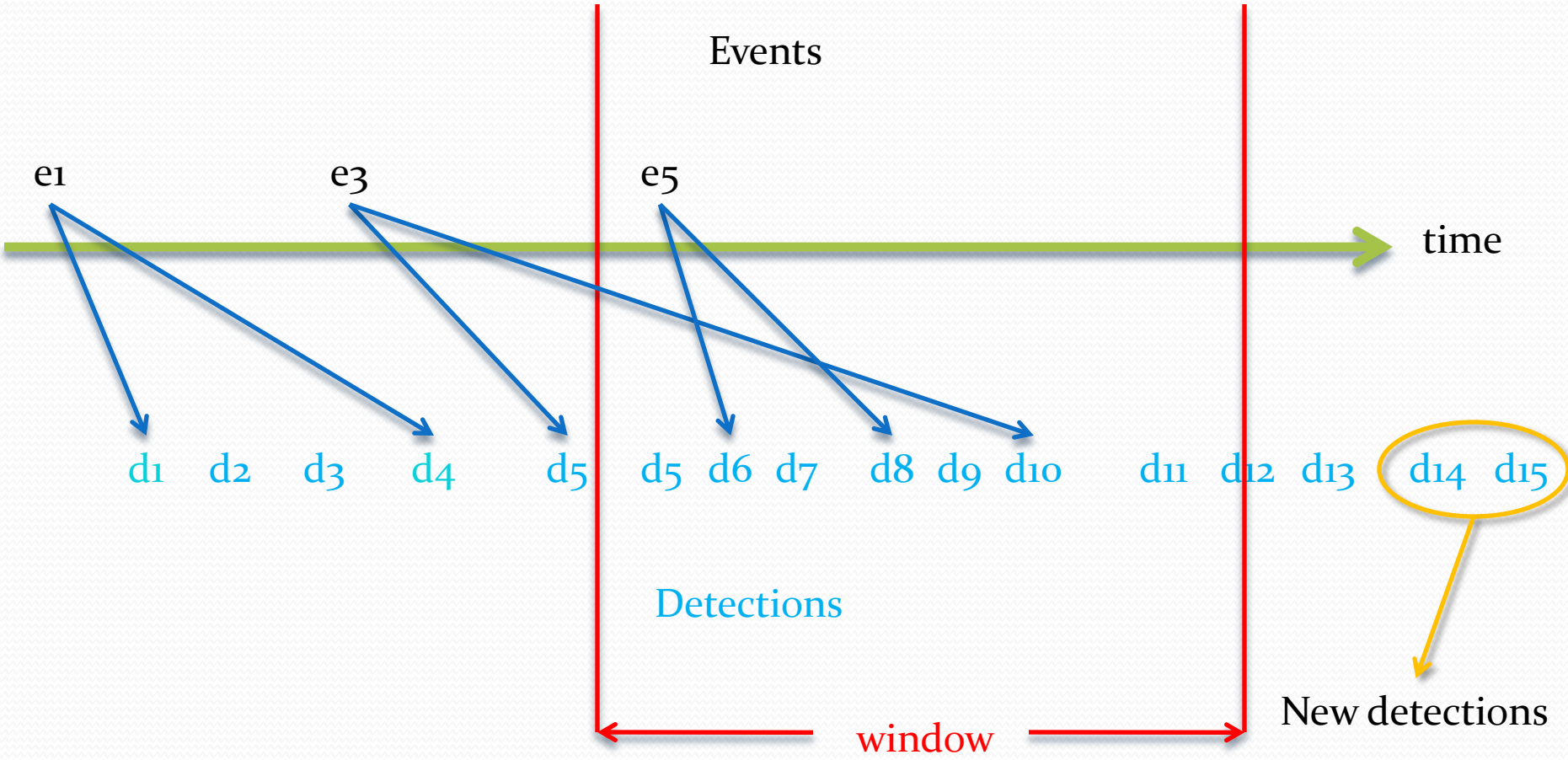
Inference : Death Move



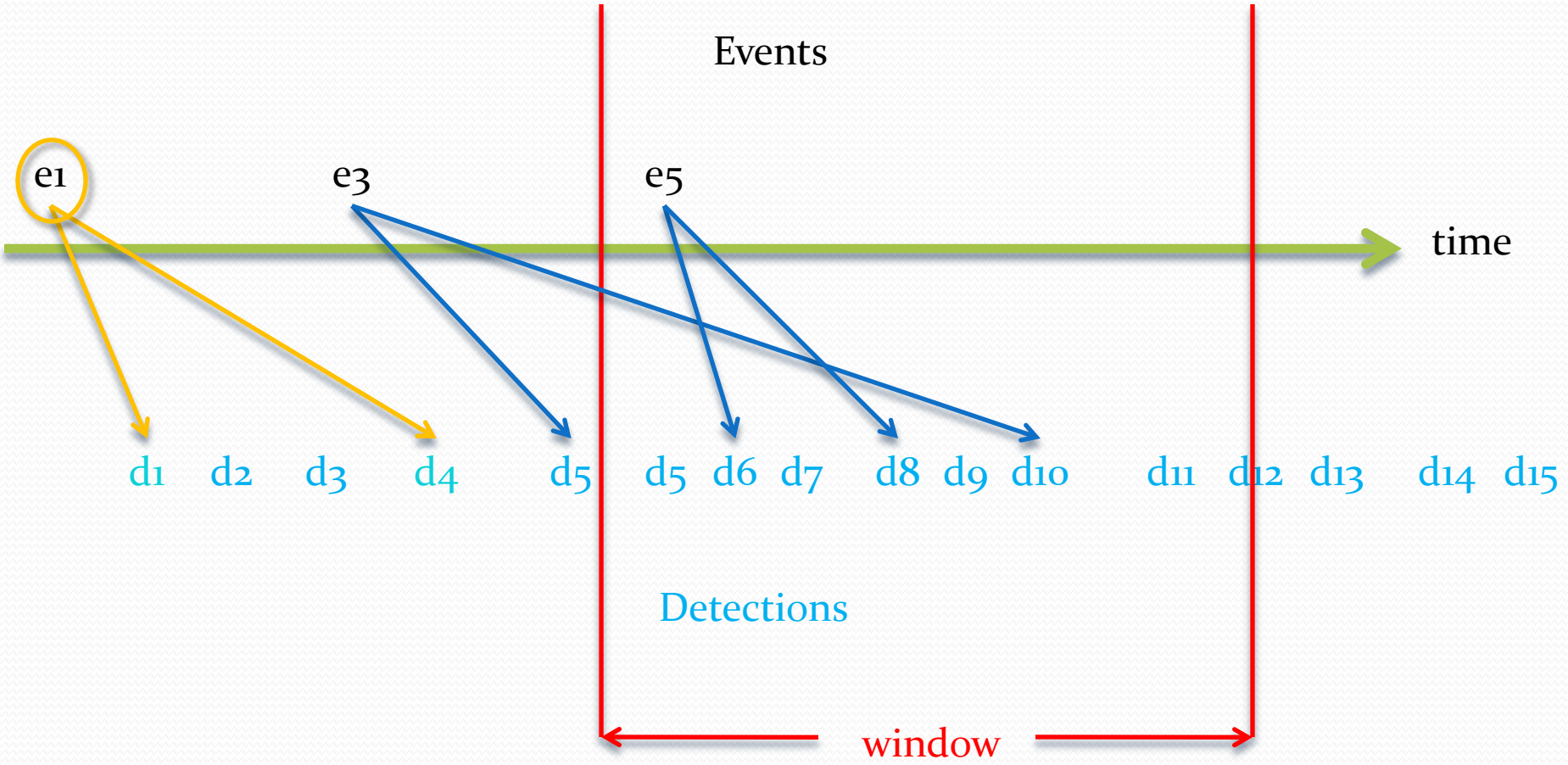
Inference : Death Move



Inference : Move Window Forward



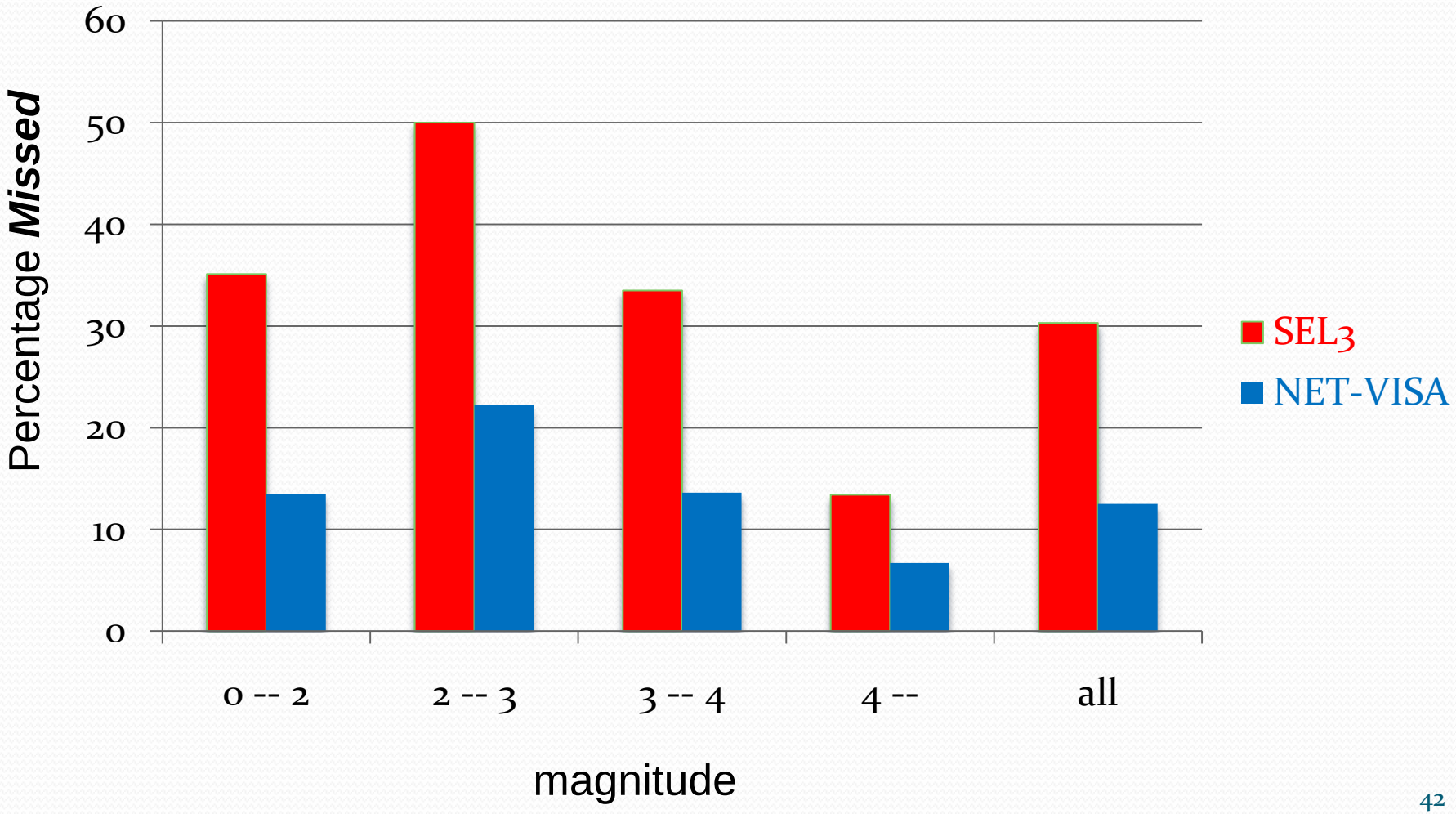
Inference : Output stable events



Overview

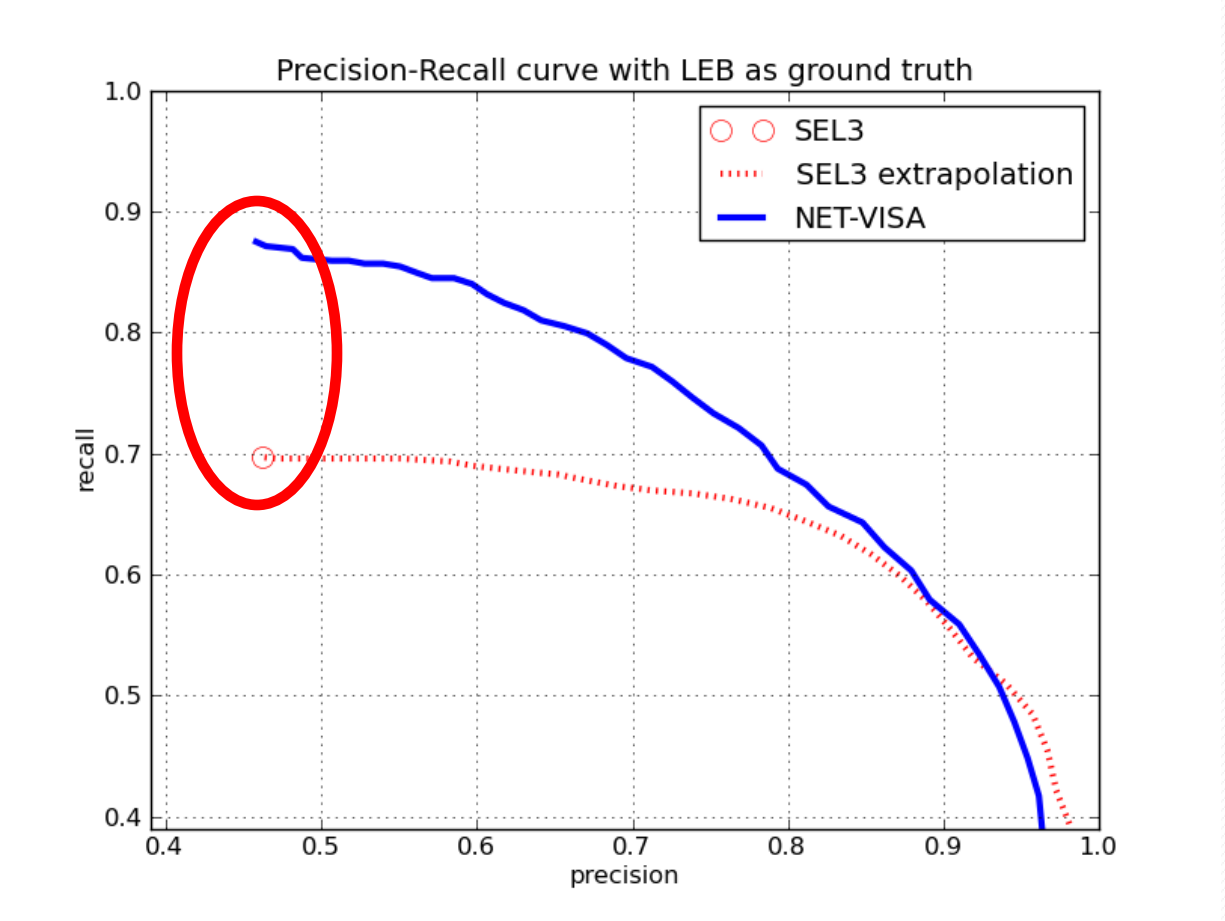
- Seismology
- Generative Model
- Inference
- Results

Percentage of Missed Events by Event Magnitude



Precision & Recall

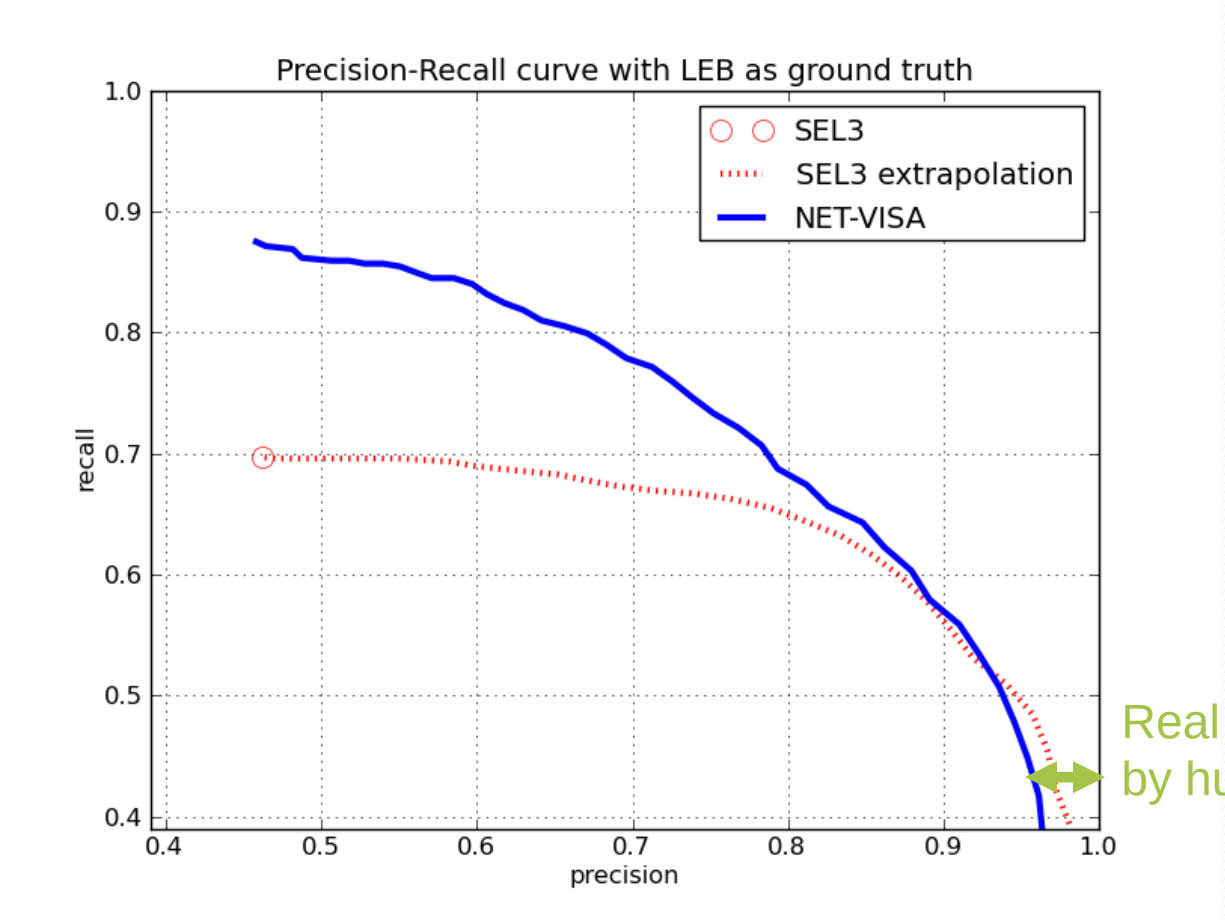
Recall = fraction of true events that are reported



Precision = fraction of reported events that are true

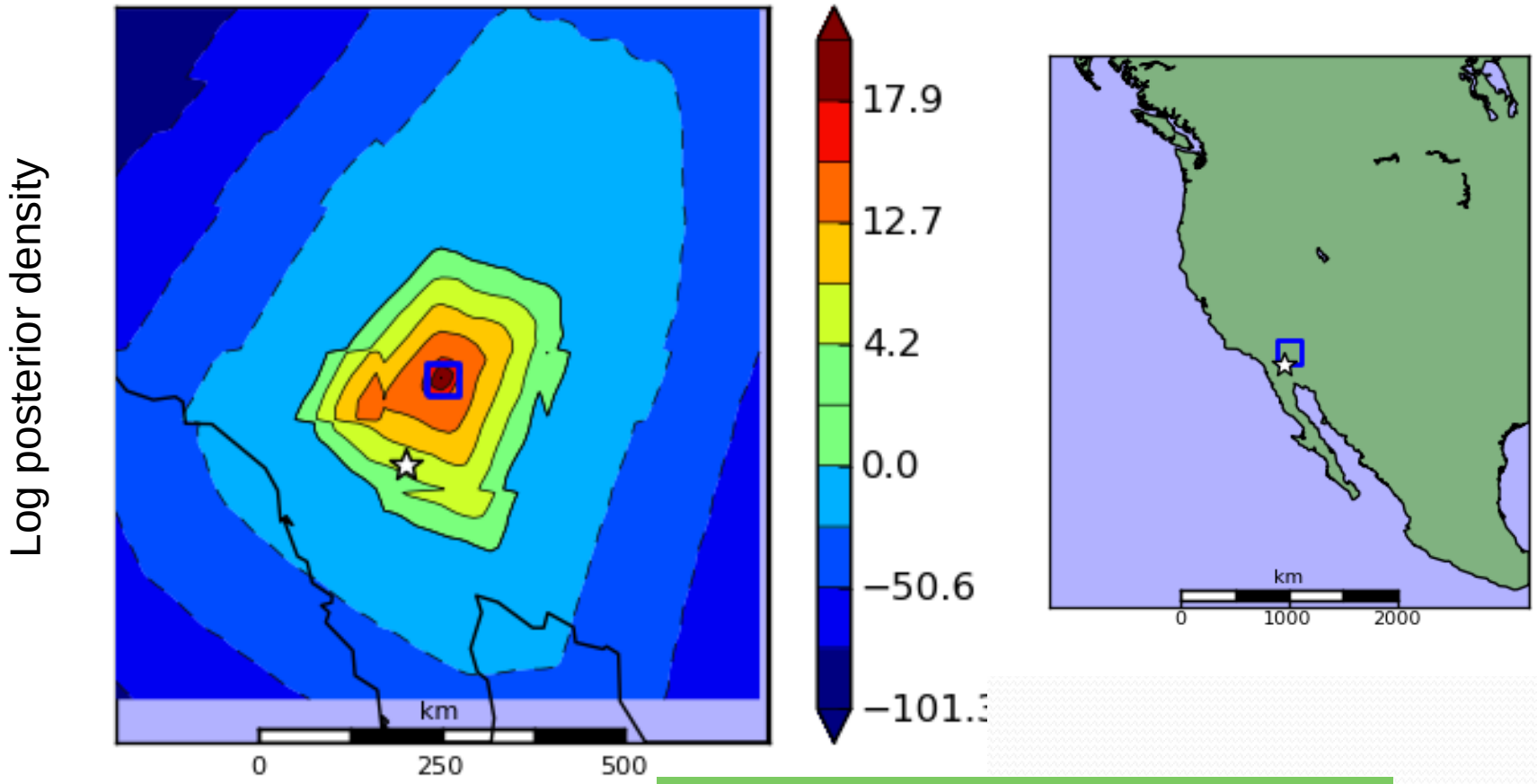
SEL3 extrapolation courtesy Mackey, Kleiner, and Jordan

Precision & Recall



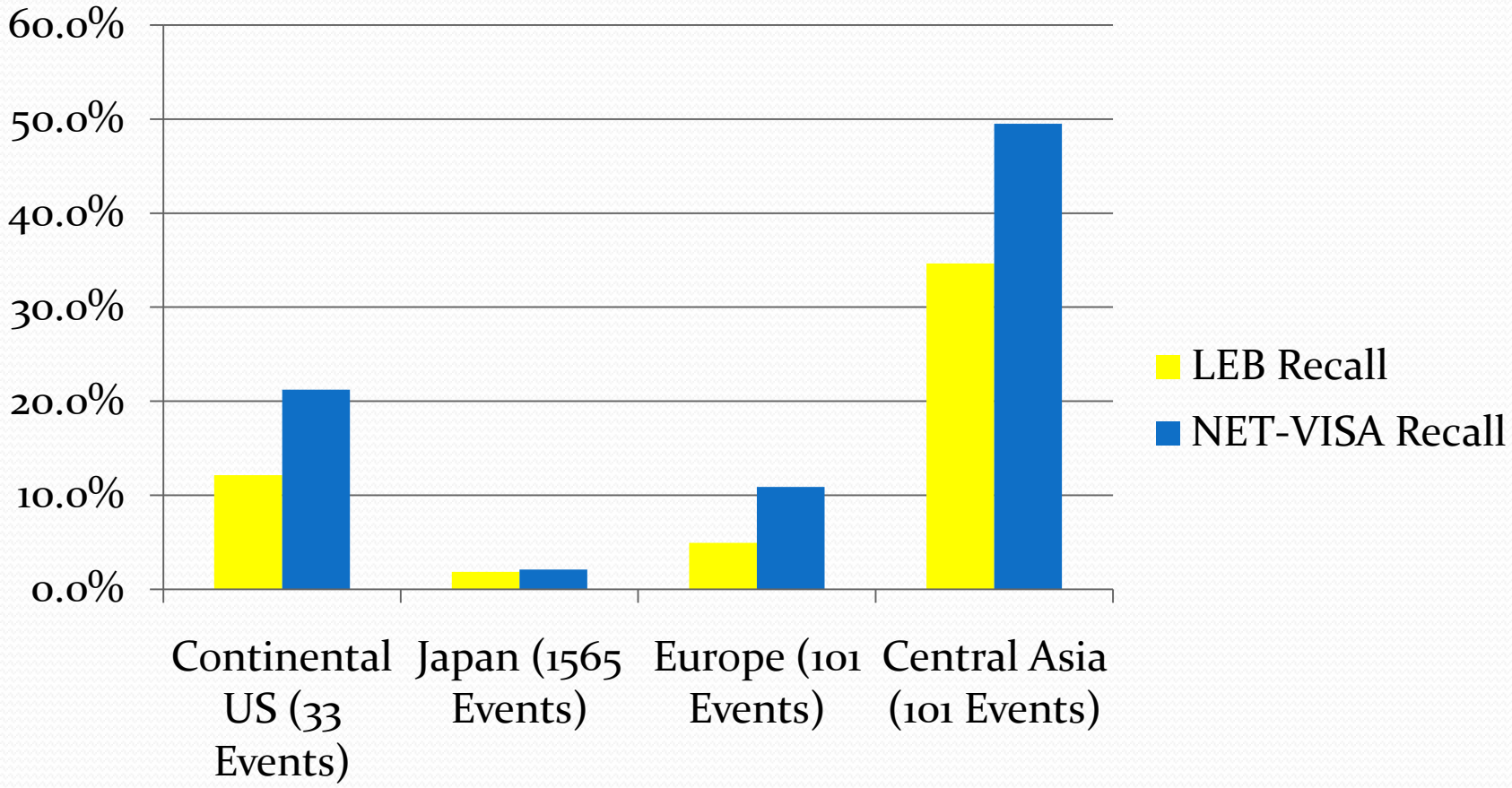
Real events missed by human analysts

NEIC Event missed by LEB

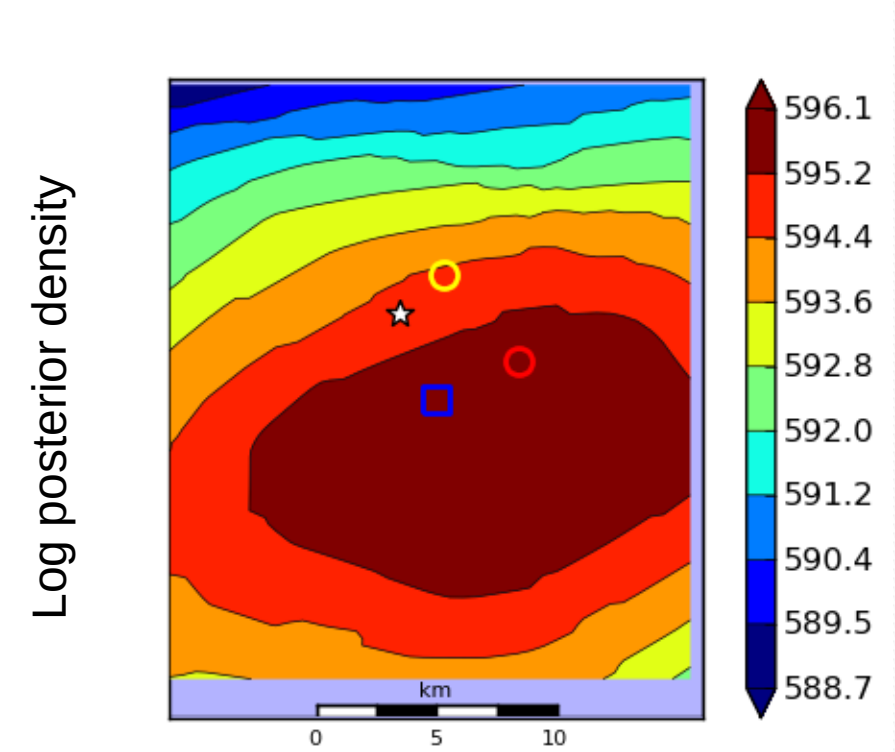


NEIC – white, and NET-VISA – blue.

Recall on Regional Networks



North Korean Explosion - 5/25/09



SEL3 – 39 stations
NET-VISA – 53 stations

NEIC – white, LEB – yellow, SEL3 – red, and NET-VISA – blue.

Conclusion

- Vertically integrated probability model of a complex, real-world process
- Combines domain knowledge with learning
- 2.5 x reduction in missed events compared to current UN system
- Currently installed for testing on CTBTO research platform
- Anticipated deployment in 2012
- Next steps:
  - Implement within general-purpose BLOG engine
  - Extend generative model to waveform level (SIG-VISA)